ENGIN. LIBRARY
TA
N58

Journal of Research of the National Bureau of Standards Vol. 47, No. 1, July 1951 Research Paper 2220

The Role of Current Distribution in Cathodic Protection

Homer D. Holler

The paper outlines a procedure for determining current distribution over an electrode surface, as required in cathodic protection or in electroplating when the electrode potential bears a known relation to current density; and shows the relation of current distribution to resistance of current path and counter electromotive force. The primary current distribution (without polarization) is also computed. A method is suggested for determining the electrode potential-current density relation over an extensive surface in a uniform medium. In a nonuniform medium, the determination of current density by measurement of electrode potential becomes complex. In such a case the potential criterion of cathodic protection must be relied upon.

I. Introduction

The electrical requirement for complete cathodic protection of a metal from corrosion was demonstrated years ago by Mears and Brown [1].[1] Their work established a criterion based on the equalization of surface potentials, which is accomplished by polarizing the cathodic areas of the metal until their potentials become equal to the "open-circuit" potential of the anodic areas. As a result, the current leaving the anodic areas, and consequently the equivalent rate of corrosion, is reduced to zero. The mechanism of the process is based on increasing the polarization of the cathodic areas by the application of external current to those areas.

Let us consider the potential [2] relations, when a voltage, E, is applied to a galvanic couple in which the potential of the cathodic area is e_c and that of the anodic area is e_a. If E is gradually increased from zero, current I will flow first to the cathodic area when $E>E_G$, where E_G is the potential of the couple and lies between e_a and e_c. That is,

$$E_G=e_a-i_0r_a$$

$$=e_c+i_0r_c,$$

where

i_0=the current circulating within the couple before E is applied
r_a=resistance of the anodic path
r_c=resistance of the cathodic path.

Only when $E>e_a$ does current begin flowing to the initially anodic area, and this occurs when $e_c+I_cr_c=e_a$, where I_c is the total cathodic current; and e_a becomes equal to E_a the open-circuit potential.

As the potential relations within a galvanic couple are such that applied current flows to the cathodic areas as required, it may at first appear that there is no problem of current distribution in the application of cathodic protection. That is, the problem seems to be automatically solved. However, let us consider the typical condition when galvanic corrosion is most insidious, that is, localized pitting. Then, the anodic area is negligible, and the problem becomes a practical one of obtaining current distribution over the cathodic area, which is approximately the whole area. In any study of the factors that determine current distribution over an extensive cathode area, the geometry and dimensions of the metal structure and its surrounding medium are determining factors. It is the purpose of this paper to analyze the potential-current relations that control this current distribution, as related to cathodic protection.

II. Dimensions and Current Distribution

In cathodic protection, we deal with cells of all shapes and sizes. It has been theoretically demonstrated [2] that the over-all potential-current relations in a large cell may not be truly represented by the results obtained in a model of a very much smaller size. This is true because the potential differences within a cell, through which current is flowing, consist of two kinds of components. One comprises electromotive forces which, for given current densities, are independent of the size of the cell; and the other includes those potential differences resulting from currents flowing through resistances which, for a given resistivity, are determined by dimensions. The smaller cell may therefore not be a true electrical model of the larger one for a given current density unless the resistivity of its electrolyte is so adjusted that the resistive components in the two cells are equal.

If the required adjustment of resistance is impracticable, the potential-current relations found in a small model may not be applicable directly to a large cell. A procedure for determining these relations by direct measurement in the cell therefore seems desirable. This is particularly true in the case of electrical circuits such as those involved in electrolysis mitigation and cathodic protection of underground pipelines where the resistive components may not be realizable on a laboratory scale. Here we have large cathodic areas, and the volume

[1] Figures in brackets indicate the literature references at the end of this paper.
[2] The term "potential" used herein really means a potential difference; and if a current is flowing, a polarized potential is understood.

938851—51——1

of electrolyte is unlimited. In the application of cathodic protection, the position of the anode, through which the external current is supplied, is of considerable practical importance. It is the purpose of this paper to outline a method tried in the laboratory, which might be translated into a field procedure for determining current distribution. Field experience will then determine whether the laboratory procedure is applicable to underground conditions.

Theoretical methods of computing current attenuation along a conductor of great length frequently neglect the role of counter emf and polarization [3], and consider the resistance as the only controlling factor. On the latter assumption, the current density at different points on the cathode varies inversely as the resistance of the current paths from the anode to the respective points, in accordance with Ohm's law. However, any counter emf reduces the current in the same ratio as that of the counter emf to the applied emf at that point. Any increase in counter emf therefore tends to reduce the current at points of higher current density to a greater degree than at points of lower current density. This results in a more uniform current distribution, for a given applied voltage, than if resistance were the sole current-limiting factor. For example, if i_1 and i_2 are the currents flowing to unit areas having electrode potentials ϵ_1 and ϵ_2, respectively, then [4]

$$\frac{\partial i_1}{\partial i_2}=\frac{r_2+\frac{\partial \epsilon_2}{\partial i_2}}{r_1+\frac{\partial \epsilon_1}{\partial i_1}}, \qquad (1)$$

where r_1 and r_2 are the resistances of the paths of the currents. For a given metal and environment,

$$\frac{\partial \epsilon_2}{\partial i_2}=\frac{\partial \epsilon_1}{\partial i_1}=\frac{\partial \epsilon}{\partial i}.$$

If $\partial\epsilon/\partial i$ becomes very large as compared with the resistances, r and r_2, $\partial i_1/\partial i_2$ approaches 1. In electroplating, this phenomenon is called "throwing power", which is a function of $\partial\epsilon/\partial i$, and the geometry of the cell and may be expressed in several ways. In cathodic protection also, a high value for $\partial\epsilon/\partial i$, as compared with resistance, favors throwing power and therefore greater uniformity in current distribution.

In cells of very small dimensions, as in pits and crevices on the metallic surface, the effective resistances r and r_2 may be small; and the value of $\partial\epsilon/\partial i$ probably greater because of larger ionic concentration gradients. These conditions favor a higher throwing power than over large areas free from sharp surface irregularities.

III. Determination of Current Distribution

In any study of current distribution, it is essential that a method of determining apparent current density at any point be available. When current distribution is uniform,

$$i=\frac{I}{A},$$

where I is the total current; and A is the total area. If I and A can be measured, there is no problem in determining the apparent current density. However, when the current distribution is not uniform, the relation of current density to polarization, if there is such a relation, may be used to determine i at any point where the cathode potential ϵ_c can be measured. If the potential ϵ_c can be measured without including any resistive components, then the relation, $\epsilon_c=(f)i$, may be used to determine i at any point in cells, regardless of dimensions.

In order to demonstrate the relation of current distribution to polarization and resistance, a test cell, large enough to obtain a convenient current-density gradient, was used. It consisted of a wooden tank, 3 ft by 10 ft by 1 ft deep, entirely insulated from outside circuits (fig. 1). A ¾-in. steel tube extended the full length of the tank and a small steel anode, A, was in one corner, as indicated. The electrolyte covered the tube by several inches. The problem was to determine the current density at points 1, 2, . . ., 10 inclusive, for a given applied emf, E_t. Direct measurement of current density at any point was impracticable. However, if the relation of cathode potential ϵ_c to i is known, then i can be easily determined. In the present case, the current-density gradient along only one dimension is of interest. For this reason current density is expressed as current per unit length of tube. Variation in current density around the tube is disregarded.

In order to obtain the data for a graph showing the relation between ϵ_c and i, measurement of ϵ_c on a given length of tube, on which the apparent current density is assumed to be uniform, is necessary at different values of I. The obvious method of obtaining such uniformity is by the use of a parallel anode. This was accomplished by replacing the small steel anode, A, with a rod equal to the length of the tank. Then $i=I/l$, where l is the length of tube, uniform current distribution being assumed. For the determination of ϵ_c, a saturated calomel electrode was placed on the surface of the tube at each numerical location, and the potential difference (ϵ_c-E_s), was measured, using the circuit in figure 1, where E_s is the potential of the calomel electrode.

IV. Method of Measurement

This circuit [5] permitted measurement of the quantities (ϵ_c-E_s) and $(\epsilon_c-E_s+Ir_s)$, where r_s is the resistance between the reference electrode and the cathode surface. For a given current, I, the bridge was balanced by adjustment of X until momentary closing of key K_1, caused *no change* in the deflection of null-indicator G. Since the resistances in arms DD were equal (each 50,000 ohms), then at balance $r_s=X$. After balance, the counter emf V_c was

sted until there was *no deflection* of G. Then ircuit (1)

$$(\epsilon_c - E_s) - V_g = (I - I_D) r_s - I_D D,$$

nd in circuit (2)

$$V_g = (I - I_D) X - I_D D.$$

Subtracting the equation for (2) from that for (1), nd since $r_s = X$,

$$(\epsilon_c - E_s) = 2V_{g1},$$

where V_{g1} = voltmeter reading.

If key K_2 be opened and V_g be readjusted to a new value V_{g2} until G again reads zero, then $(\epsilon_c - E_s + Ir_s) = V_{g2}$, where V_{g2} is the voltmeter reading. This is equivalent to a potentiometer measurement. When Ir_s is negligible, use of the bridge is unnecessary, and $(\epsilon_c - E_s)$ may be observed directly as V_{g2}, for all practical purposes.

The instrumental requirements of this circuit are not stringent. As null indicator G, a General Electric galvanometer was satisfactory. In the field, a Weston Model 622 voltmeter with zero-center scale and a resistance of 200,000 ohms/v may be suitable and more convenient. A calibrated slide-wire rheostat of sufficient current-carrying capacity was suitable for balancing the bridge and reading. This circuit has been used in preliminary field tests but is not yet in suitable portable form.

V. Potential Criterion of Cathodic Protection

For each value of I, the emf $(\epsilon_c - E_s)$, was measured at each point in numerical order, and also in the reverse order, making a total of 20 readings. The results are summarized in figure 2, where two graphs represent the averages, one for tap water and one for 0.1-percent NaCl solution. The "breaks" occur at potential levels near but slightly below the open-circuit potential of iron in a saturated solution of ferrous hydroxide. At this potential, 0.813 volt, corresponding to a pH of about 9.5, the reaction $Fe \rightleftarrows Fe^{++} + 2\epsilon$ reaches equilibrium, and corrosion of iron by this process practically ceases [6]. The data in figure 2 were obtained for electrolytes that were undisturbed, except by convection, and were very probably saturated with air. An air-free environment is therefore not a requirement for obtaining the air-free potential by cathodic polarization. The reproducibility of the readings was better than was anticipated for iron in a solution initially having an undefined concentration of ferrous ions.

The protective current, I_p, indicated in the salt solution was about 2.4 ma/ft and in the tap water, 1.5 ma/ft. Thus, the corrosiveness in the more conductive salt solution was greater than in the tap water, though not in the ratio of the conductivities, which was approximately ten to one. It is apparent that this large difference in conductivity had no apparent effect on the potential at which the breaks in the polarization curves occurred

Next, the parallel anode was replaced by the point anode A, and for a given value of E_t and I, ϵ_c was observed at the numbered points along the tube. By the use of the graphs in figure 2, the corresponding values of i were then estimated for each point. In table 1 the values of ϵ_c and i are given, for each point of observation, in a 0.1-percent solution of sodium chloride for a current of 12 ma. Since the potential of the tube $= \epsilon_0$, when $i = 0$, the difference $\Delta\epsilon_c = \epsilon_c - \epsilon_0$ represents the emf of polarization, also given in table 1.

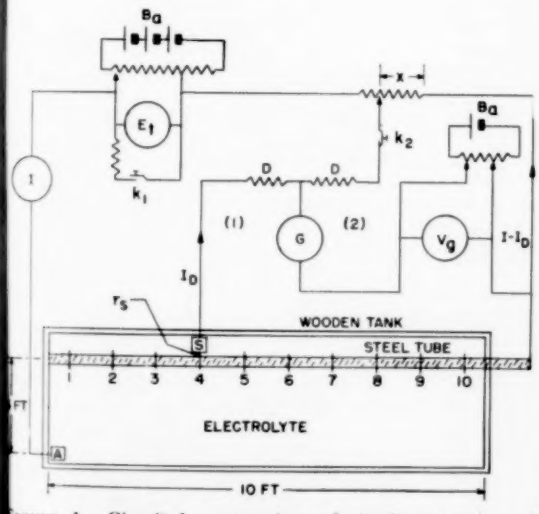

FIGURE 1. *Circuit for measuring polarization at points of different current densities along a steel tube.*

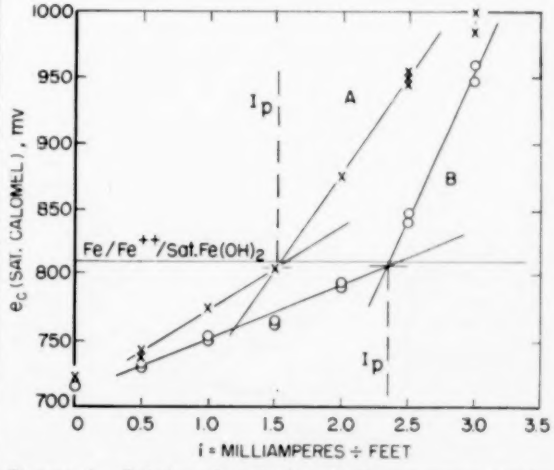

FIGURE 2. *Relation of electrode potential to current density along a steel tube.*

A, Tap water; B, 0.1-percent NaCl solution.

TABLE 1. *Observed potentials and current densities at different points; calculated resistances and current-density components*

e_c=715 mv; e_A=588 mv; E_t=1,170 mv; I=12 ma.

Point number	e_c	Δe_c	i	r	E_t/r	$\frac{e_c-e_A}{r}$	$\frac{\Delta e_c}{r}$
	mv	*mv*	*ma/ft*	*Ohms*	*ma/ft*	*ma/ft*	*ma/ft*
1	958	−243	3.00	268	4.37	−1.37	−0.90
2	928	−213	2.85	294	3.98	−1.13	−.70
3	861	−146	2.55	353	3.31	−0.76	−.40
4	803	−88	2.15	445	2.63	−.48	−.19
5	774	−59	1.30	762	1.54	−.24	−.07
6	764	−49	1.05	952	1.23	−.18	−.05
7	749	−34	0.60	1695	0.69	−.09	−.02
8	746	−31	.60	1695	.69	−.09	−.02
9	748	−33	.60	1695	.69	−.09	−.02
10	756	−41	.75	1351	.87	−.12	−.03

VI. Calculation of Resistance of Current Path

Let us now analyze the relation of current-density attenuation to the resistance of the current paths from the anode to successive unit lengths of tube. For a given current I, the polarized anode potential=e_A. Then, since the applied emf=E_t, and the external resistance was negligible,

$$\frac{E_t-(e_c-e_A)}{i}=r, \qquad (2)$$

where r=the resistance of the path of current i from anode A to a given unit length of tube. That is, $r_1, r_2, \ldots, r_{10}$ are the resistances of the paths of currents $i_1, i_2, \ldots, i_{10}$, respectively. Using eq 2, the values of r were computed from the corresponding values of e_c and i, where E_t=1,170 mv, and e_A=588 mv. These results are also summarized in table 1.

VII. Polarization

Rearranging eq 2, we obtain

$$i=\frac{E_t}{r}-\frac{(e_c-e_A)}{r}, \qquad (3)$$

where i consists of two components in opposition. One, E_t/r, with E_t constant depends only on r; and the other $(e_c-e_A)/r$, with e_A constant, depends on r and also upon e_c. The expression E_t/r is the primary current distribution representing the current density that would be obtained if there were no counter emf. For example, if the cell were entirely filled uniformly with a conductor of the first class, or if it had identical nonpolarizable electrodes, like copper in copper sulphate solution at low current densities, $(e_c-e_A)=0$. Thus E_t/r_t is the largest current density attainable for a given value of E_t and r.

If $e_c>e_A$ and e_c is nonpolarizable, having a constant value e_0, then we have a constant counter emf (e_0-e_A), as for example, in a lead-acid storage battery on charge at low current density.

If $e_c>e_A$ and e_c is polarizable, then we have the present case, where (e_c-e_A) represents the counter emf in which $e_c=e_0+\Delta e_c$. The amount of polarization may be expressed as Δe_c in millivolts, or as $\Delta e_c/r$ in milliamperes. Both indices in table 1 show that the role of polarization in determining current-density distribution diminishes at the lower current densities. At the minimum current density, resistance is practically the sole determining factor, as indicated in figure 3.

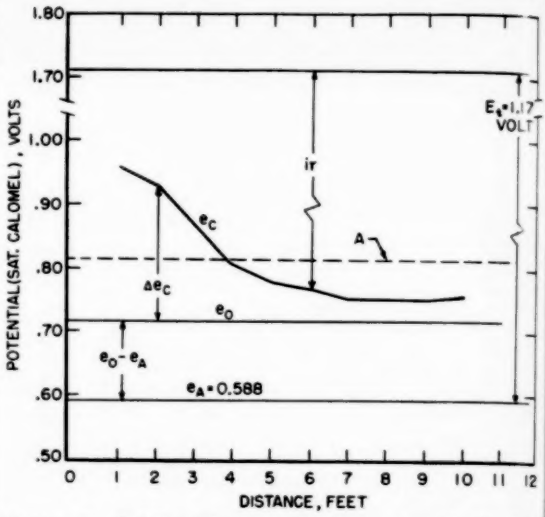

FIGURE 3. *Relation between emf and ir components of polarization along a steel tube.*

A, Open-circuit potential of iron, 0.834 volt.

The current-distribution curves for salt water (fig. 4) were essentially the same as those obtained in the tap water (fig. 5). For example, cathodic protection defined as that value of e_c, equal to or above 0.813 v, extended approximately to the 4-ft location in each electrolyte; and as shown in (fig. 2), a lower current density was required for protection in the tap water than in the salt solution.

VIII. Location of Reference Electrode

When the tip of the reference electrode was in contact with the steel tube surface, there was a small resistance r_s (fig. 1) between the electrode and the surface, which increased with current density (fig. 6). On a copper tube, this resistance was greater by several-fold and may have been due to liberated hydrogen, or to a film resulting from increased alkalinity at the surface.

When the reference electrode was located 18 in. from the tube surface, the observed readings were averages of the true potentials for a range of points extending over a considerable length of the tube. For example, in figure 6, the potential (800 mv) observed at locations 1 and 2, with the electrode 18 in. away was approximately the average potential over a 5-ft length as measured with the reference electrode directly on the surface. The significance of this observation is that, when the potential and

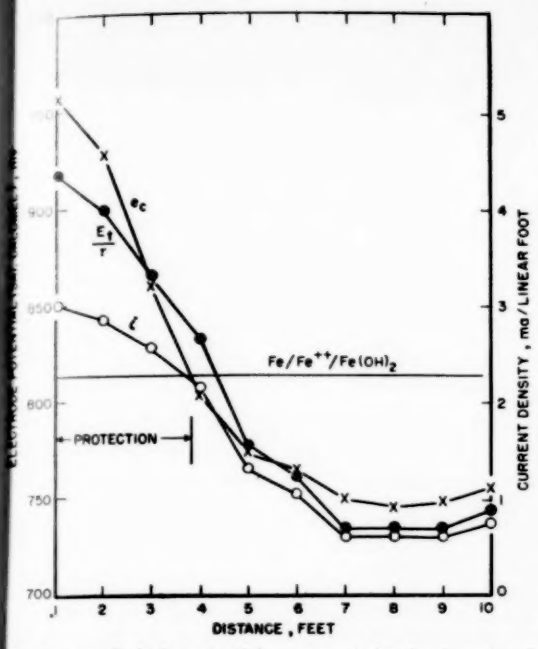

FIGURE 4. *Cathodic potential e_c, current density i, and primary current distribution E_t/r along a steel tube in* NaCl *solution.*

0.5-Percent NaCl solution; I=12 ma.

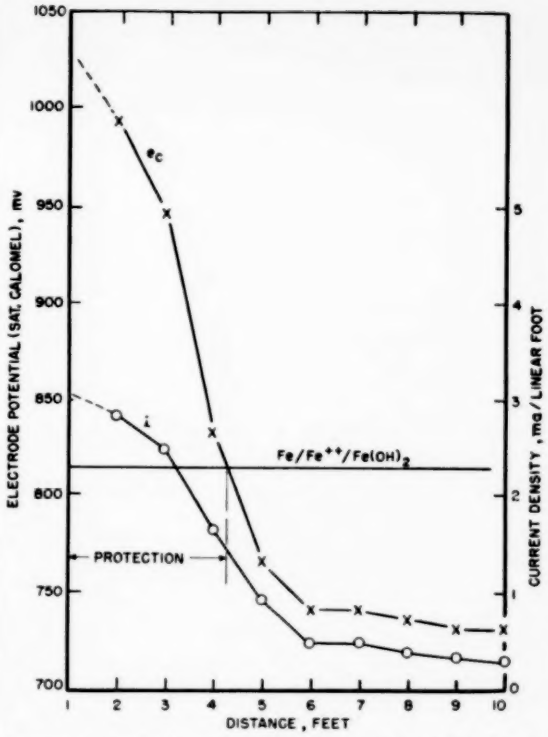

FIGURE 5. *Cathodic potential e_c and current density i along a steel tube in Washington tap water.*

Tap water; I=12 ma.

current-density gradients along the tube are steep, the reference electrode must be as near as possible to the surface in order that the observed value approach the magnitude of e_c at a given point. At best, the observed value of e_c is an average potential over an area that increases with the distance of the reference electrode from the surface. The purpose of reducing this distance to a minimum is not that of reducing the effect of resistance, since this can be balanced out by the bridge measurement, but in order to reduce the observed area to a minimum. Obviously, if the cathode potential is uniform, the requirement of proximity is unnecessary. A more remote location may then be used, the distance from the surface being greater, in proportion to the area of uniformity.

A frequent subterfuge of locating the reference electrode behind the electrode under investigation, or at some other point of minimum current density, may neither eliminate the *ir* component in the observed potential nor indicate the value of e_c at the front surface of the electrode, which is of most importance. The reference electrode should therefore be placed at the exact point, where the value of e_c is required, with the least possible disturbance of the lines of current flow. This requires a reference electrode with the smallest exploring tip that is practicable.

IX. Suggested Field Procedure for Determining Current Distribution

In estimating the current distribution over a very extended cathode surface, such as a pipe-line or a very large plane surface, the cathodic polarization-current density relation may be very useful. For determining this relation, a procedure, based on the guard-ring principle [7] in electrical measurements, suggests itself. For example in figure 7, let a parallel three-section line anode be laid as near the pipe surface as possible without distrubing the surface environment. The two end-sections will supply current to unknown lengths of pipe over which the current-density varies from a maximum to practically zero at some unknown distance away. By thus eliminating the "end-effect" current from the meter reading, we know the length of pipe which is receiving the measured current I, being supplied by the middle section of the anode.

The degree of uniformity of current density within the length l, of course, may not be predictable. If the values of e_c observed at numerous points within length l and the resistivity of the surrounding medium along the length l are fairly uniform, then assuming

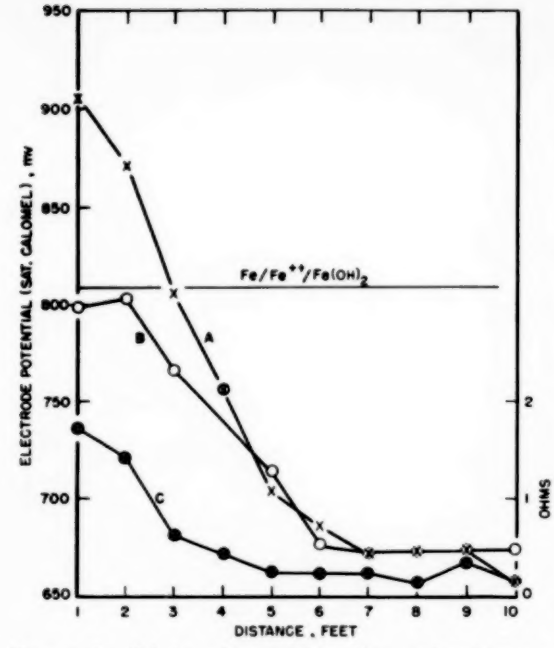

FIGURE 6. *Relation of potential readings and resistance r_s along a steel tube, for different positions of the reference electrode.*

Steel tube in 0.1-percent NaCl solution; 12 ma; A, reference electrode on surface; B, reference electrode 18 in. from surface; C, resistance r_s with reference electrode on surface.

that e_c is a function of i the current distribution should be uniform. If it is, then a clearly defined potential-current density relation similar to that in figure 2 should be obtainable by the technique outlined in figure 7. Such a relation may then be used to determine current distribution beyond the ends of length l in a uniform environment.

In the case of a large plane surface, l in Figure 7 may represent the diameter of a disk within a ring having a slightly larger inside diameter, both being in a plane near and parallel to the surface. The measured current, I, will flow to an area approximately equal to that of the disk, the current density i being $4I/\pi l^2$. Then the relation between e_c and i may be established within the disk area by the bridge-method used in figure 1. If $e_c=(f)i$, the current-density gradient outside the disk area may be determined by measuring e_c at points at increasing distances away. This is very much simpler than the classical but laborious procedure of plotting the equi-potential surfaces around the electrode and then graphically determining the current density lines perpendicular to them.

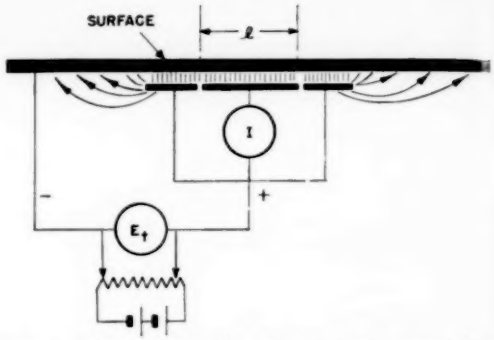

FIGURE 7. *Suggested procedure for measuring current flowing to a known length of pipe-line.*

X. Conclusions

Using a reference electrode and a bridge method, which eliminates ir components from the measurement, the relation between current density and the polarized potential of iron was determined in tap water and in a 0.1-percent salt solution. By the application of this relation to a large cathode receiving current from a point anode, the current density at different points on the cathode was obtained from the electrode potentials measured at the same points. The resistances of the current paths from the anode to cathode were computed; and the role of resistance and counter emf in current distribution, thereby demonstrated.

By this procedure, current distribution may be determined, regardless of the dimensions of the cell, if the reference electrode is located very near the cathode surface. If the surface potential is uniform, the reference electrode may be located at some remote point. If the surface potential is not uniform, the reference electrode at a remote point will indicate an average potential over a given area.

When the polarized potential is a known function of current density, the primary current distribution may be computed from the data obtainable by the above procedure. By using the guard-ring principle, the potential-current density relation, and current distribution may be determined over very extensive electrode areas in a uniform environment.

XI. References

[1] R. B. Mears and R. H. Brown, Trans. Electrochem. Soc. **74,** 519 (1938).
[2] J. N. Agar and T. P. Hoar, Discussions on electrode processes, Faraday Soc. No. 1, 158 (1947).
[3] C. Kasper, Trans. Electrochem. Soc. **78,** 131 (1940).
[4] H. D. Holler, J. Research NBS **45,** 373 (1950) RP2148.
[5] H. D. Holler, J. Electrochem. Soc. **97,** 271 (1950).
[6] H. D. Holler, Ind. Eng. Chem. **21,** 750 (1929).
[7] J. J. Thomson, Recent researches in electricity and magnetism, p. 233 (Clarendon Press, Oxford, Eng., 1893).

WASHINGTON, December 13, 1950.

...l of Research of the National Bureau of Standards Vol. 47, No. 1, July 1951 Research Paper 2221

Multiplets and Terms in Technetium Spectra

William F. Meggers

Multiplets and terms of the quartet, sextet, and octet systems are reported for the first spectrum of Tc. The combinations of 20 terms composed of 63 atomic energy levels account for 200 Tc I lines and 66 percent of the total observed intensity summed for 1,300 lines.

The ground state of neutral Tc atoms is represented by $4d^5\,5s^2\,{}^6S_{2½}$, and the first metastable state by $4d^6 5s\,{}^6D_{4½}$, only 2573 cm^{-1} higher. A Rydberg series of ^8S terms yields absolute term values from which a Ritz-corrected limit of 60070 cm^{-1}, or ionization potential of 7.45 electron volts, is derived.

In the second spectrum of Tc the principal S and P terms of the quintet and septet systems are identified and fixed relative to each other. Evidence is presented that $4d^6\,{}^5D$ represents the normal state of Tc^+ ions, but to confirm this further observations of the Tc II spectrum are necessary.

I. Introduction

In 1922 Catalán [1] [1] discovered, in the spectrum of manganese, that certain groups of lines having the same character are related by very exact numerical separations, and he suggested that the word "multiplet" be used to denote such groups. Since that time the search for multiplets in all atomic spectra has been the chief objective of spectrum analysis. This search, aided by the quantum theory, has resulted in present-day information concerning spectral terms, atomic structures, ionization potentials, and many other properties of atoms, of ions, and of atomic nuclei. The purpose of this paper is to report on information of this type derived from preliminary analyses of the first two spectra of technetium, the most recently discovered chemical homolog of manganese.

This chemical element was predicted in 1869, by Mendeléeff who found that when the then-known chemical elements were arranged in order of increasing atomic weights, their maximum chemical valencies exhibited periodicities, with gaps suggesting two unknown homologs of manganese. In 1925 Noddack and Tacke [2] announced the discovery of both of these elements and proposed for them the names "masurium" and "rhenium." Rhenium soon became available in abundance, and in 1931 Meggers published a preliminary description of its arc spectrum, also reporting multiplets, atomic energy levels, and the ionization potential for Re I. The separation of masurium from natural minerals has never been repeated or confirmed, but the element in question has recently been produced artificially either by transmutation of molybdenum [3] or by fission of uranium [4]. This artificial element was named [5] technetium (symbol Tc). A preliminary description of the arc and spark spectra of Tc by Meggers and Scribner [6] constitutes the material in which the first multiplets and spectral terms of Tc were found.

The basic spectral data consist of measured wavelengths and estimated relative intensities of more than 2,200 radiations characteristic of Tc atoms and ions. These spectral lines are divided into two classes, one (Tc I) characteristic of neutral Tc atoms, and the other (Tc II) of singly ionized atoms. A considerable number of Tc lines were observed to be complex (*c*) with hyperfine structure, and many were hazy (*h*) in the spark spectrograms. Other criteria, such as reversibility, pressure displacement, temperature classification, and Zeeman effect, which are extremely helpful in the analyses of complex spectra, are still wholly lacking for Tc lines. In the absence of these aids it might appear presumptuous to attempt the analyses of Tc spectra, but fortunately these spectra are so well-observed and orthodox that the principal multiplets could be identified without difficulty. The first guiding principles in the search for Tc multiplets were the interval and intensity rules, assuming that *LS*-coupling controls the optical electrons. The second clues to Tc multiplets came from comparisons of analogous Mn and Tc spectra as regards term intervals, term magnitudes, and electron configurations. Under these conditions a reliable list of lines, wavelengths, and intensities will yield atomic information, provided that the observed spectral region and range are right. This statement is illustrated by the atomic information derived from the analyses of Tc I and Tc II spectra.

[1] Figures in brackets indicate the literature references at the end of this paper.

II. First Spectrum of Technetium, Tc I

1. Tc I Multiplets

The observed data of the Tc I spectrum comprise about 1,300 lines with wavelengths ranging from 2324.57 to 8829.80 A, and estimated intensities ranging from 1 to 500. Vacuum wave numbers corresponding to the measured wavelengths of Tc I lines were obtained from Kayser's Tabelle der Schwingungszahlen [7]. Like other first spectra of metals, the Tc I spectrum is weak in the ultraviolet, strong in the short-wave visible and adjacent ultraviolet, and lacks lines of high intensity in the near infrared. My search for multiplets in this spectrum began with the identification of a group of three lines ($a\,{}^6S_{2½}-z\,{}^6P^\circ_{3½,\ 2½,\ 1½}$) which are outstanding

in the analogous Mn I (4030.76, 4033.07, 4034.49 A) and Re I (3451.88, 3466.47, 3464.72 A) spectra.

Naively, I interpolated between Mn I and Re I, and looked for the analogous Tc I lines between 3700 and 3800 A, but failed to find them there. The most intense Tc I lines have wavelengths 4297.06, 4262.26, 4238.19 A. The wave number separations (189.9 and 133.2 cm^{-1}) of these three lines are almost exactly proportional to 7 and 5, which characterize a good ^6P term. The same separations were next found among seven weak lines (4948 to 5139 A), but the significance of these was not recognized until recurring differences (302.1 and 449.7 cm^{-1}) found between wave numbers of strong lines (3500 to 4200 A) were seen to be identical with the cross differences of the seven weak lines. This led at once to the identification and fixation of the metastable five-level term, $a\ ^6D$, and many multiplets that involve this term. Surplus $a\ ^6D$ separations among other Tc I lines were regarded as belonging to intersystem combinations, which gave the clue to the fixation and identification of a metastable four-level term $a\ ^4D$, and multiplets that involve this term.

Finally, a unique pair of visible Tc I lines (5924.57 and 6085.22 A), exhibiting the widest hyperfine structure, were assumed to represent the transition $a^6S_{2\frac{1}{2}} - z\ ^6P^\circ_{3\frac{1}{2},2\frac{1}{2}}$.

This assumption was confirmed by the establishment of $e\ ^6S$ and $e\ ^6D$, which give with $z\ ^6P^\circ$ the strongest lines in the octet system. The principal multiplets of the Tc I spectrum are displayed in tables 1, 2, and 3, where the term symbols and values (relative to $a\ ^6S_{2\frac{1}{2}} = 0.0$) are shown at the top and left. Differences of these level values are entered in the body of the table where the levels intersect. These numbers are the actual wave numbers of all lines permitted by the selection rule $\Delta J = 0, \pm 1$. The observed wavelengths and intensities (parentheses) appear directly above the wave numbers. The average difference between computed and observed wave numbers is ± 0.06 cm^{-1}, which proves that the measured wavelengths of these Tc lines are rarely in error by more than 0.01 A in relative value. The total number of permitted lines in tables 1, 2, 3 is 240, whereas the number of observed lines is 200. In a few cases the unobserved Tc I lines are obscured by Tc II, by impurity, or by strong copper lines, but in most cases the absence of predicted combinations is explained by the fact that they are either weak intersystem transitions or double-electron jumps. One line 4145.0 A is doubly classified; the calculated components are separated by 1 cm^{-1} but were not resolved because of unequal intensities.

TABLE 1. *Multiplets in the* Tc I *spectrum*

	$a\ ^6S_{2\frac{1}{2}}$ 0. 0	$a\ ^6D_{4\frac{1}{2}}$ 2572. 9	$a\ ^6D_{3\frac{1}{2}}$ 3250. 9	$a\ ^6D_{2\frac{1}{2}}$ 3700. 6	$a\ ^6D_{1\frac{1}{2}}$ 4002. 7	$a\ ^6D_{0\frac{1}{2}}$ 4178. 8	$e\ ^6S_{2\frac{1}{2}}$ 39599. 0
$z\ ^6P^\circ_{3\frac{1}{2}}$	4297.06(500)	4831.33(3)	4995.00(15)	5109.79(8)			6120.67(150)
23265.4	23265.2	20692.5	20014.5	19564.8			16333.6
$z\ ^6P^\circ_{2\frac{1}{2}}$	4262.26(400)		4948.06(8)	5060.69(30)	5139.25(10)		6192.65(100)
23455.3	23455.2		20204.3	19754.7	19452.7		16143.7
$z\ ^6P^\circ_{1\frac{1}{2}}$	4238.19(300)			5026.79(5)	5104.33(30)	5150.58(10)	6244.17(50)
23588.5	23588.4			19887.9	19585.8	19409.9	16010.5
$z\ ^6D^\circ_{4\frac{1}{2}}$		4031.63(300)	4145.02(80)				
27369.8		24796.9	24118.6				
$z\ ^6D^\circ_{3\frac{1}{2}}$		3984.97(150)	4095.68(150)	4172.52(80)			
27660.1	27660.1	25087.2	24409.1	23959.5			
$z\ ^6D^\circ_{2\frac{1}{2}}$			4049.10(150)	4124.22(60)	4176.28(100)		
27940.8	27940.8		24689.9	24240.2	23938.0		
$z\ ^6D^\circ_{1\frac{1}{2}}$	3551.20(1)			4088.70(100)	4139.85(5)	4170.28(50)	
28151.4	28151.5			24450.8	24148.7	23972.6	
$z\ ^6D^\circ_{0\frac{1}{2}}$					4115.08(70)	4145.02(80)	
28296.7					24294.0	24118.6	
$z\ ^6F^\circ_{5\frac{1}{2}}$		3636.10(300)					
30967.1		27494.2					
$z\ ^6F^\circ_{4\frac{1}{2}}$		3627.39(25)	3718.88(200)				
30133.1		27560.2	26882.2				
$z\ ^6F^\circ_{3\frac{1}{2}}$		3594.93(2)	3684.78(100)	3746.87(150)			
30382.0	30382.0	27809.1	27131.1	26681.4			
$z\ ^6F^\circ_{2\frac{1}{2}}$			3664.94(3)	3726.36(100)	3768.80(100)		
30528.8	30528.8		27277.9	26828.3	26526.1		
$z\ ^6F^\circ_{1\frac{1}{2}}$				3712.28(20)	3754.41(50)	3779.40(50)	
30630.5	30630.5			26930.0	26627.8	26451.7	
$z\ ^6F^\circ_{0\frac{1}{2}}$					3746.18(15)	3771.06(60)	
30689.1					26686.3	26510.3	
$x\ ^6P^\circ_{3\frac{1}{2}}$	3182.38(30)	3466.29(200)	3549.74(150)	3607.35(50)			
31414.0	31414.0	28841.1	28163.1	27713.3			
$x\ ^6P^\circ_{2\frac{1}{2}}$	3183.12(30)		3550.66(100)	3608.30(100)	3648.07(80)		
31406.7	31406.7		28155.8	27706.1	27404.0		
$x\ ^6P^\circ_{1\frac{1}{2}}$	3173.30(40)			3595.70(50)	3635.18(80)	3658.62(80)	
31503.8	31503.9			27803.1	27501.1	27325.0	

TABLE 1. *Multiplets in the* Tc I *spectrum*—Continued

	$a\ ^6S_{2\frac{1}{2}}$ 0.0	$a\ ^6D_{4\frac{1}{2}}$ 2572.9	$a\ ^6D_{3\frac{1}{2}}$ 3250.9	$a\ ^6D_{2\frac{1}{2}}$ 3700.6	$a\ ^6D_{1\frac{1}{2}}$ 4002.7	$a\ ^6D_{0\frac{1}{2}}$ 4178.8	$e\ ^6S_{2\frac{1}{2}}$ 39599.0
$z\ ^4P_{2\frac{1}{2}}$	3131.26(15)		3486.25(50)	3541.78(50)	3580.08(20)		
31927.0	31926.8		28676.0	28226.4	27924.4		
$z\ ^4P_{1\frac{1}{2}}$	3099.11(6)			3500.72(20)	3538.70(10)	3560.34(12)	
32258.0	32258.0			28557.4	28255.5	28079.2	
$w\ ^4P_{3\frac{1}{2}}$	2614.24(50)	2802.84(20)	2857.15(12)	2894.34(6)			
38240.6	38240.6	35667.6	34989.7	34540.1			
$w\ ^4P_{2\frac{1}{2}}$	2615.88(40)		2859.13(40)	2896.36(12)	2921.94(6)		
38216.6	38216.7		34965.4	34516.0	34213.8		
$w\ ^4P_{1\frac{1}{2}}$	2608.88(30)			2887.76(15)	2913.17(12)	2928.20(10)	
38319.4	38319.2			34618.8	34316.8	34140.7	
$y\ ^4D_{4\frac{1}{2}}$		2730.53(20)	2782.08(25)				
39184.8		36612.1	35933.8				
$y\ ^4D_{3\frac{1}{2}}$		2762.16(7)	2814.88(15)	2850.97(8)			
38766.0	38766.0	36192.9	35515.1	35065.5			
$y\ ^4D_{2\frac{1}{2}}$	2589.88(7)		2828.07(10)	2864.51(10)			
38600.3	38600.3		35349.4	34899.8	34597.6		
$y\ ^4D_{1\frac{1}{2}}$				2762.36(10)	2785.60(15)		
39891.0	39891.0			36190.3	35888.3	35712.2	
$y\ ^4D_{0\frac{1}{2}}$					2794.55(4)	2808.38(15)	
39776.1					35773.4	35597.3	
$z\ ^4F_{4\frac{1}{2}}$		3502.73(50)	3587.96(80)				
31114.0		28541.0	27863.1				
$z\ ^4F_{3\frac{1}{2}}$	3163.17(1)	3443.49(3)	3525.84(30)	3582.66(40)			
31604.9	31604.7	29032.0	28354.0	27904.3			
$z\ ^4F_{2\frac{1}{2}}$	3122.66(6)		3475.60(20)		3568.87(20)		
32014.7	32014.7		28763.8	28314.1	28012.1		
$z\ ^4F_{1\frac{1}{2}}$	3099.54(4)			3501.26(4)	3538.70(10)	3560.90(1)	
32253.5	32253.5			28553.0	28250.9	28074.6	
$z\ ^4D_{3\frac{1}{2}}$		3327.12(3)	3303.91(2)	3456.86(6)			
32620.3	32620.3	30047.4	29369.6	28919.7			
$z\ ^4D_{2\frac{1}{2}}$			3350.86(2)		3437.45(3)		
33085.6	33085.6		29834.5	29385.0	29083.0		
$z\ ^4D_{1\frac{1}{2}}$					3407.27(3)	3227.87(1)	
33343.1	33343.1			29642.5	29340.6	29164.3	
$z\ ^4D_{0\frac{1}{2}}$							
33480.2					29477.5	29301.4	
$y\ ^4P_{2\frac{1}{2}}$	2896.36(12)		3197.53(3)	3244.19(10)	3276.30(2)		
34516.0	34516.0		32265.1	30815.5	30513.5		
$y\ ^4P_{1\frac{1}{2}}$	2860.98(2)						
34942.4	34942.8			31241.8	30939.7	30763.6	
$y\ ^4P_{0\frac{1}{2}}$							
35276.2					31273.5	31097.4	

TABLE 2. *Multiplets in the* Tc I *spectrum*

	$a\ ^6S_{2½}$ 0.0	$a\ ^4D_{3½}$ 10516.5	$a\ ^4D_{2½}$ 11063.0	$a\ ^4D_{1½}$ 11578.5	$a\ ^4D_{½½}$ 11890.7
$z\ ^4P_{2½}$		5411. 92(2)	5576. 77(2)	5741. 92(1)	
28989. 4	28989. 4	18472. 6	17926. 6	17411. 0	
$z\ ^4P_{1½}$			5439. 48(1)	5596. 53(1)	5696. 03(2)
29441. 9	29441. 9		18379. 0	17863. 3	17551. 2
$z\ ^4P_{½½}$				5455. 94(3)	5550. 51(3)
29902. 1				18323. 6	18011. 4
$z\ ^4F_{4½}$		4853. 57(400)			
31114. 0		20597. 6			
$z\ ^4F_{3½}$	3163. 17(1)	4740. 59(200)	4866. 71(250)		
31604. 9	31604. 7	21088. 5	20542. 0		
$z\ ^4F_{2½}$	3122. 66(6)	4650. 23(5)	4771. 53(100)	4891. 89(150)	
32014. 8	32014. 7	21498. 3	20951. 8	20436. 3	
$z\ ^4F_{1½}$	3099. 54(4)		4717. 76(10)	4835. 40(30)	4909. 55(100)
32253. 5	32253. 5		21190. 6	20675. 0	20362. 8
$z\ ^4D_{3½}$		4522. 85(200)	4637. 51(90)		
32620. 3	32620. 3	22103. 8	21557. 3		
$z\ ^4D_{2½}$		4429. 60(30)	4539. 54(100)	4648. 34(70)	
33085. 6	33085. 6	22569. 1	22022. 5	21507. 1	
$z\ ^4D_{1½}$			4487. 05(40)	4493. 36(20)	4660. 21(30)
33343. 1	33343. 1		22280. 1	21764. 5	21452. 3
$z\ ^4D_{½½}$				4564. 57(30)	4630. 57(20)
33480. 2				21901. 7	21589. 6
$y\ ^4P_{2½}$	2896. 36(12)	4165. 62(80)	4262. 67(100)	4358. 49(3)	
34516. 0	34516. 0	23999. 3	23452. 9	22937. 5	
$y\ ^4P_{1½}$	2860. 98(2)		4186. 52(50)	4278. 91(30)	4336. 87(10)
34942. 4	34942. 8		23879. 4	23363. 9	23051. 6
$y\ ^4P_{½½}$				4218. 61(20)	4274. 96(5)
35276. 2				23697. 8	23385. 5
$z\ ^6D_{4½}$		5931. 93(60)			
27369. 8		16853. 3			
$z\ ^6D_{3½}$		5831. 45(30)	6023. 49(4)		
27660. 1	27660. 1	17143. 6	16597. 1		
$z\ ^6D_{2½}$		5737. 52(1)	5923. 35(10)		
27940. 8	27940. 8	17424. 3	16877. 7	16362. 3	
$z\ ^6D_{1½}$	3551. 20(1)			6032. 34(6)	
28151. 4	28151. 5		17088. 4	16572. 7	16260. 7
$z\ ^6D_{½½}$					6093. 71(2)
28296. 7				16718. 2	16405. 8
$z\ ^6F_{5½}$		5096. 27(200)			
30133. 1		19616. 7			
$z\ ^6F_{4½}$		5032. 44(3)	5174. 79(100)		
30382. 0	30382. 0	19865. 5	19319. 1		
$z\ ^6F_{3½}$			5135. 78(1)	527 [illegible]9)	
30528. 8	30528. 8	20012. 3	19465. 8	189[illegible]	
$z\ ^6F_{2½}$					5334. 79(15)
30630. 5	30630. 5		19567. 5	19052. 0	18739. 7
$z\ ^6F_{1½}$				5231. 22(1)	5318. 16(1)
30689. 1				19110. 7	18798. 3
$y\ ^6P_{3½}$	3131. 26(15)	4669. 30(100)	4791. 62(4)	4913. 01(15)	
31927. 0	31926. 8	21410. 5	20864. 0	20348. 5	
$y\ ^6P_{2½}$	3099. 11(6)		4716. 76(1)	4834. 36(30)	4908. 50(40)
32258. 0	32258. 0		21195. 1	20679. 5	20367. 2
$w\ ^6P_{3½}$	2614. 24(50)		3678. 41(2)		
38240. 6	38240. 6	27724. 1	27177. 9		
$w\ ^6P_{2½}$	2615. 88(40)	3609. 06(1)	3681. 71(1)	3752. 93(1)	
38216. 6	38216. 7	27700. 2	27153. 6	26638. 4	
$w\ ^6P_{1½}$	2608. 88(30)				
38319. 4	38319. 2		27256. 4	26740. 9	26428. 7

TABLE 3. *Multiplets in the* Tc I *spectrum*

	$z\,^5P_{4\frac{1}{2}}$ *17523.0*	$z\,^5P_{3\frac{1}{2}}$ *16874.5*	$z\,^5P_{2\frac{1}{2}}$ *16428.8*
$a\,^5S_{2\frac{1}{2}}$		5924. 57(100)	6085. 22(100)
0. 0		16874.5	16428. 7
$a\,^5D_{4\frac{1}{2}}$	6687. 03(3)	6990. 22(1)	
2572. 9	14950. 2	14301. 8	
$a\,^5D_{3\frac{1}{2}}$		7338. 18(3)	
3250. 9	14272. 1	13623. 6	13177. 9
$a\,^5D_{2\frac{1}{2}}$			7854. 35(1)
3700. 6		13173. 9	12728. 3
$a\,^5D_{1\frac{1}{2}}$			8045. 30(2)
4002. 7			12426. 2
$c\,^5S_{2\frac{1}{2}}$	4976. 34(200)	4820. 75(150)	4719. 28(100)
37612. 5	20089. 5	20737. 9	21183. 8
$c\,^5D_{4\frac{1}{2}}$	3724. 41(50)		
44365. 3	26842. 3		
$c\,^5D_{3\frac{1}{2}}$	3726. 19(10)	3638. 24(40)	
44352. 5	26829. 5	27478. 0	
$c\,^5D_{2\frac{1}{2}}$	3727. 37(3)	3639. 41(30)	3581. 27(20)
44343. 9	26821. 0	27469. 2	27915. 1
$c\,^5D_{1\frac{1}{2}}$		3640. 23(10)	3582. 10(15)
44337. 5		27463. 0	27908. 7
$c\,^5D_{0\frac{1}{2}}$			3582. 66(40)
44333. 1			27904. 3
$f\,^5S_{2\frac{1}{2}}$	3119. 18(3)	3057. 34(2)	3016. 31(1)
49573. 3	32050. 4	32698. 7	33143. 5

Although the 200 classified lines presented here comprise only 15 percent of the total number of Tc I lines, they account for 66 percent of the total intensity, summed for 1,300 observed lines.

2. Tc I Spectral Terms

The Tc I multiplets in tables 1, 2, 3 result from combinations of 20 spectral terms embracing 63 atomic energy levels that are arrayed in table 4 alongside the analogous terms and levels belonging to the Mn I spectrum [8]. The electron configurations in column 1 apply to Mn I; for Tc I the principal quantum numbers must be increased by one unit. Naturally, the intervals, between levels of complex terms are larger in Tc than in Mn, but the algebraic signs, with few exceptions, are the same. A plus sign means that the term is regular, minus means inverted. The Mn I terms are arranged in order of increasing magnitude of levels having the largest J but the analogus Tc I terms depart from this order because the sixth d-type electron is more firmly bound in Tc than in Mn. Thus, in Mn I $3d^5\,4s\,{}^6D_{4\frac{1}{2}}$ is 17052.3 cm^{-1} above the ground state $3d^5\,4s^2\,{}^6S_{2\frac{1}{2}}$, whereas in Tc I $4d^5\,5s\,{}^6D_{4\frac{1}{2}}$ is only 2572.9 cm^{-1} above zero. Similarly, the associated $^4D_{3\frac{1}{2}}$ levels are, respectively, 23296.7 and 10516.5 cm^{-1} above zero. It appears that terms from configurations containing six d-electrons are displaced about 13,000 cm^{-1}, whereas those from configurations with five d-electrons are displaced only about 2,000 cm^{-1}. These displacements are shown in the last column of table 4; they served as a guide in assigning Tc I terms to the electron configurations in the first column. After the relative energies of $d^n s$ and $d^{n-1}\,s^2$ had been determined for Tc as described above, it was called to my attention that an approximate value of this difference was predicted in 1928 by Gibbs and White [9].

TABLE 4. *Spectral terms of* Mn I *and* Tc I

Electron configuration	Term symbol	J	Mn I		Tc I		Mn I—Tc I
			Level	Interval	Level	Interval	
$3d^5\,4s(^7S^3S)4s$	$a\,^6S$	$2\frac{1}{2}$	0. 0		0. 0		0
$3d^6(a\,^3D)4s$	$a\,^6D$	$4\frac{1}{2}$	17052. 3		2572. 9		14479
		$3\frac{1}{2}$	17282. 0	−229. 7	3250. 9	−678. 0	14031
		$2\frac{1}{2}$	17451. 5	−169. 5	3700. 6	−449. 7	13751
		$1\frac{1}{2}$	17568. 5	−117. 0	4002. 7	−302. 1	13566
		$0\frac{1}{2}$	17637. 2	−68. 7	4178. 8	−176. 1	13458
$3d^5\,4s(a\,^7S)4p$	$z\,^6P°$	$4\frac{1}{2}$	*18705. 4*		*17523. 0*		1182
		$3\frac{1}{2}$	*18531. 7*	173. 7	*16874. 5*	648. 5	1657
		$2\frac{1}{2}$	*18402. 5*	129. 2	*16428. 8*	445. 7	1974
$3d^6(a\,^5D)4s$	$a\,^4D$	$3\frac{1}{2}$	23296. 7		10516. 5		12780
		$2\frac{1}{2}$	23549. 2	−252. 5	11063. 0	−546. 5	12486
		$1\frac{1}{2}$	23719. 5	−170. 3	11578. 5	−515. 5	12141
		$0\frac{1}{2}$	23818. 8	−99. 3	11890. 7	−312. 2	11928
$3d^5\,4s(a\,^7S)4p$	$z\,^6P°$	$3\frac{1}{2}$	*24802. 2*		*23265. 4*		1537
		$2\frac{1}{2}$	*24788. 0*	14. 2	*23455. 3*	−189. 9	1333
		$1\frac{1}{2}$	*24779. 3*	8. 7	*23588. 5*	−133. 2	1191

TABLE 4. *Spectral terms of* Mn I *and* Tc I—Continued

Electron configuration	Term symbol	J	Mn I		Tc I		Mn I—Tc I
			Level	Interval	Level	Interval	
$3d^5\ 4s(a\ ^5S)4p$	$z\ ^4P^\circ$	$2\frac{1}{2}$	*31001. 1*		*28989. 4*		2012
		$1\frac{1}{2}$	*31076. 4*	−75. 3	*29441. 9*	−452. 5	1634
		$0\frac{1}{2}$	*31124. 9*	−48. 5	*29902. 1*	−460. 2	1223
$3d^5\ 4s(a\ ^5S)4p$	$y\ ^6P^\circ$	$3\frac{1}{2}$	*35769. 9*				
		$2\frac{1}{2}$	*35725. 8*	44. 1	*31927. 0*		3799
		$1\frac{1}{2}$	*35690. 0*	35. 8	*32258. 0*	−331. 1	3432
$3d^5\ 4s(a\ ^7S)5s$	$e\ ^8S$	$3\frac{1}{2}$	39431. 3		37612. 5		1819
$3d^5\ 4s(a\ ^7S)5s$	$e\ ^8S$	$2\frac{1}{2}$	41403. 9		39599. 0		1805
$3d^6(a\ ^5D)4p$	$z\ ^6D^\circ$	$4\frac{1}{2}$	*41789. 4*		*27369. 8*		14420
		$3\frac{1}{2}$	*41932. 6*	−143. 2	*27660. 1*	−290. 3	14272
		$2\frac{1}{2}$	*42053. 7*	−121. 1	*27940. 8*	−280. 7	14113
		$1\frac{1}{2}$	*42143. 5*	−89. 8	*28151. 4*	−210. 6	13992
		$0\frac{1}{2}$	*42198. 6*	−55. 1	*28296. 7*	−145. 3	13902
$3d^6(a\ ^5D)4p$	$z\ ^6F^\circ$	$5\frac{1}{2}$	*43314. 2*		*30067. 1*		13247
		$4\frac{1}{2}$	*43428. 6*	−114. 4	*30133. 1*	−66. 0	13296
		$3\frac{1}{2}$	*43524. 1*	−95. 5	*30382. 0*	−248. 9	13142
		$2\frac{1}{2}$	*43595. 5*	−71. 4	*30528. 8*	−146. 8	13067
		$1\frac{1}{2}$	*43644. 4*	−48. 9	*30630. 5*	−101. 8	13014
		$0\frac{1}{2}$	*43673. 0*	−28. 6	*30689. 1*	−58. 6	12984
$3d^6(a\ ^5D)4p$	$z\ ^4F^\circ$	$4\frac{1}{2}$	*44288. 8*		*31114. 0*		13175
		$3\frac{1}{2}$	*44523. 5*	−234. 7	*31604. 9*	−490. 9	12919
		$2\frac{1}{2}$	*44696. 3*	−172. 8	*32014. 8*	−409. 9	12682
		$1\frac{1}{2}$	*44814. 6*	−118. 3	*32253. 5*	−238. 7	12561
$3d^6(a\ ^5D)4p$	$x\ ^6P^\circ$	$3\frac{1}{2}$	*44993. 9*		*31414. 0*		13580
		$2\frac{1}{2}$	*45156. 1*	−162. 2	*31406. 7*	7. 3	13749
		$1\frac{1}{2}$	*45259. 2*	−103. 1	*31503. 8*	−97. 1	13755
$3d^6(a\ ^5D)4p$	$z\ ^4D^\circ$	$3\frac{1}{2}$	*45754. 2*		*22620. 3*		13134
		$2\frac{1}{2}$	*45940. 9*	−186. 7	*33085. 6*	−465. 3	12855
		$1\frac{1}{2}$	*46083. 9*	−143. 0	*33343. 1*	−257. 5	12741
		$0\frac{1}{2}$	*46169. 9*	−86. 0	*33480. 2*	−137. 1	12690
$3d^5\ 4s(a\ ^7S)4d$	$e\ ^8D$	$5\frac{1}{2}$	46712. 6		44365. 3		2347
		$4\frac{1}{2}$	46710. 2	2. 4	44352. 5	12. 8	2358
		$3\frac{1}{2}$	46708. 3	1. 9	44343. 9	8. 6	2364
		$2\frac{1}{2}$	46707. 0	1. 3	44337. 5	6. 4	2370
		$1\frac{1}{2}$	46706. 1	0. 9	44333. 1	4. 4	2373
$3d^6\ 4s(a\ ^5D)4p$	$y\ ^4P^\circ$	$2\frac{1}{2}$	*46901. 1*		*34516. 0*		12385
		$1\frac{1}{2}$	*47154. 5*	−253. 4	*34942. 4*	−426. 4	12212
		$0\frac{1}{2}$	*47299. 3*	−144. 8	*35276. 2*	−333. 8	12023
$3d^5\ 4s(a\ ^7S)5p$	$w\ ^6P^\circ$	$3\frac{1}{2}$	*47387. 6*		*38240. 6*		9147
		$2\frac{1}{2}$	*47659. 5*	−271. 9	*38216. 6*	24. 0	9443
		$1\frac{1}{2}$	*47782. 4*	−122. 9	*38319. 4*	−102. 8	9463
$3d^5\ 4s(a\ ^5P)4p$	$y\ ^6D^\circ$	$4\frac{1}{2}$	*47903. 8*		*39184. 8*		8719
		$3\frac{1}{2}$	*47774. 5*	129. 3	*38766. 0*	416. 8	9009
		$2\frac{1}{2}$	*47754. 0*	20. 5	*38600. 3*	165. 7	9154
		$1\frac{1}{2}$	*47466. 7*	287. 3	*39891. 0*	−1290. 7	7576
		$0\frac{1}{2}$	*47452. 2*	14. 5	*39776. 1*	114. 9	7676
$3d^5\ 4s(a\ ^7S)6s$	$f\ ^8S$	$3\frac{1}{2}$	50157. 6		49573. 3		584
$3d^5\ 4s(a\ ^7S)6s$	$g\ ^8S$	$2\frac{1}{2}$	50904. 6		49717. 3		1187

Departures from the Landé interval rule occur in Tc spectra, but they are occasionally much greater in Tc I than in Mn I. For example, the interval between the two largest J values of $z\,^6F^\circ$ is abnormally small in Tc I, and $x\,^6P^\circ$ is only partially inverted in Tc I. Despite careful search the $J=3\frac{1}{2}$ level of $y\,^6P^\circ$ could not be found. The intervals in $y\,^6D^\circ$ are extremely irregular in both spectra. On the whole, both term intervals and line intensities indicate that, to a large degree, the optical electrons of Tc atoms obey LS-coupling. It is planned to test this by observing the Zeeman effect, and incidentally to check the proposed designations of Tc atomic energy levels.

3. First Ionization Potential of Tc

If two or more spectral terms form a Rydberg series, that series may be extrapolated to a limiting wave number that represents the energy of binding of the last bound electron. In complex atoms, such as those under discussion, the number of excited states is relatively large in successive shells, and the transitions from high terms in series either give faint lines or are not observed. However, as long ago as 1894, Kayser and Runge [10] detected the first spectral series in Mn I. That series, now interpreted as $z\,^8P^\circ - n\,^8S$, permitted a calculation of the absolute value of $z\,^8P^\circ$, and then by addition of the intersystem combination $a\,^6S - z\,^8P^\circ$ the value of the term $a\,^6S$ representing the normal state of the atom was arrived at. A similar procedure is now possible with Tc I because two multiplets (of three lines each) have been found and interpreted as combinations of $z\,^8P^\circ$ with successive series terms $e\,^8S$ and $f\,^8S$. These combinations, $z\,^8P^\circ - n\,^8S$, as well as the intersystem lines, $a\,^6S - z\,^8P^\circ$, are shown in table 3. A Rydberg interpolation table applied to the series lines yields the following approximate values for the $z\,^8P^\circ$ levels of Tc I:

$$z\,^8P^\circ_{4\frac{1}{2}} = 42766$$
$$z\,^8P^\circ_{3\frac{1}{2}} = 43415$$
$$z\,^8P^\circ_{2\frac{1}{2}} = 43860$$

Adding the wave numbers of the resonance lines, we obtain the following absolute values for the $a^6S_{2\frac{1}{2}}$ term:

$$a\,^6S_{2\frac{1}{2}} = z\,^8P^\circ_{3\frac{1}{2}} + (a\,^6S_{2\frac{1}{2}} - z\,^8P^\circ_{3\frac{1}{2}}) = 43415 + 16875 = 60290 \text{ cm}^{-1},$$

$$a\,^6S_{2\frac{1}{2}} = z\,^8P^\circ_{2\frac{1}{2}} + (a\,^6S_{2\frac{1}{2}} - z\,^8P^\circ_{2\frac{1}{2}}) = 43860 + 16429 = 60289 \text{ cm}^{-1}.$$

Assuming that 1 electron volt $= 8067.5$ cm^{-1}, this value of the deepest term in the Tc I spectrum corresponds to an ionization potential of 7.47 electron volts.

Experience tells us that extrapolation of a two-term Rydberg series generally leads to a slightly too-large value for the limit, which is more accurately given by a Ritz formula of the type $\nu = L - R/(n+\mu+\alpha T)^2$. In the spectral series for 37Rb, 46Pd, and 47Ag the Ritz correction α has an average value of 2.0×10^{-6}. Assuming that this correction is valid also for 43Tc, the absolute value of the $a\,^6S_{2\frac{1}{2}}$ term of Tc I is found to be 60070 cm^{-1} and the corresponding ionization potential is $60070/8067.5 = 7.45$ electron volts.

III. Second Spectrum of Technetium, Tc II

1. Tc II Multiplets

Within the wavelength limits 2261.30 A and 4434.96 A about 800 spectral lines have been definitely identified as belonging to singly ionized Tc atoms, the Tc II spectrum. The short-wave limit was imposed by the smallness of the Tc sample and by the declining efficiency of photographic plates and spectrograph employed in recording the spark spectrum. The paucity of enhanced Tc II lines in the visible spectrum is responsible for the long-wave limit of observation. The measured wavelengths of Tc II lines were converted to vacuum wave numbers with Kayser's [7] Tabelle der Schwingungszahlen.

Among the observed Tc II lines there are two ultraviolet multiplets, each of three intense lines, which were immediately recognized as the combinations $a\,^7S_3 - z\,^7P^\circ_{2,3,4}$ (2647.02, 2610.00, 2543.24 A) and $a\,^5S_2 - z\,^5P^\circ_{1,2,3}$ (3195.21, 3212.01, 3237.02 A). In the Mn II spectrum $a\,^7S_3$ is the ground state, and 9473 cm^{-1} lower than $a\,^5S_2$. Assuming that a similar state of affairs could be expected in the Tc II spectrum, combinations between the quintet and septet terms were sought in order to fix their relative values. Fortunately, these intersystem combinations were unmistakeably represented by two outstanding lines (3892.14 and 3975.02 A), at the short-wave edge of the visible spectrum, and by the two most prominent lines (2298.10 and 2285.47 A) very near the ultraviolet limit of observation. These Tc II multiplets are presented in table 5, which is constructed like tables 1, 2, 3. The Tc II terms revealed by these

TABLE 5. *Multiplets in the* Tc II *spectrum*

	$a\,^7S_3$ 0.0	$a\,^5S_2$ 12617.1
$z\,^7P^\circ_4$	2543. 24(1000)	
39308. 1	39308. 1	
$z\,^7P^\circ_3$	2610. 00(800)	3892. 14(40c)
38302. 7	38302. 7	25685. 6
$z\,^7P^\circ_2$	2647. 02(600)	3975. 02(10)
37767. 1	37767. 1	25150. 0
$z\,^5P^\circ_3$	2298. 10(10)	3237. 02(500c)
43500. 8	43500. 8	30883. 7
$z\,^5P^\circ_2$	2285. 47(3)	3212. 01(300c)
43741. 3	43741.2	31124. 2
$z\,^5P^\circ_1$		3195. 21(200c)
43904. 9		31287. 8

TABLE 6. *Spectral terms of* Mn II *and* Te II

Electron configuration	Term symbol	J	Mn II		Te II		Mn II—Te II
			Level	Interval	Level	Interval	
$3d^5(a\ ^6S)4s$	$a\ ^7S$	3	0.0		0.0+x		x
$3d^5(a\ ^6S)4s$	$a\ ^5S$	2	9472.9		12617.1+x		−3244+x
$3d^5(a\ ^6S)4p$	$z\ ^7P^\circ$	4	*38806.5*	263.5	*39308.1*+x	1005.4	−502+x
		3	*38543.0*	176.9	*38302.7*+x	535.6	+240+x
		2	*38366.1*		*37767.1*+x		599+x
$3d^5(a\ ^6S)4p$	$z\ ^5P^\circ$	3	*43370.4*	−114.1	*43500.8*+x	−240.5	−130+x
		2	*43484.5*	−72.5	*43741.3*+x	−163.5	−257+x
		1	*43557.0*		*43904.9*+x		−348+x

multiplets are displayed in table 6 alongside the analogous terms in the Mn II spectrum [8]. Attention is called to the large separation (12,617 cm^{-1}) between $a\ ^7S_3$ and $a\ ^5S_2$, and to the fact that the intervals in the $z\ ^5P^\circ$ term are almost exactly in the ratio 3 to 2 as required by the Landé interval rule.

In the Mn II spectrum the second metastable state is $3d^6\ a\ ^5D$, which combines with $3d^5\ 4p\ z\ ^5P^\circ$ to produce a prominent multiplet in the ultraviolet (3441.98 to 3497.53 A). The analogous multiplet of Te II lines was sought with the aid of the $z\ ^5P^\circ$ intervals, 240.5 and 163.6 cm^{-1}. In view of the stronger binding of d-type electrons in the Te I spectrum, this Te II multiplet must be expected to lie farther in the ultraviolet. Failure to find it among the observed lines forced the conclusion that in Te II the term $a\ ^5D$ has lower energy than $a\ ^7S$, and that the multiplet under discussion will therefore have wavelengths outside the present limit of observation of Te spectra. Until this multiplet is observed it is impossible to give the correct values of Te II terms in table 6 relative to the normal state $a\ ^5D=0.0$; this uncertainty is expressed by x. Further progress in the analysis and interpretation of the second spectrum of technetium would be greatly facilitated by extending the observations to both shorter and longer waves than the present limits.

IV. References

[1] M. A. Catalán, Phil. Trans. Roy. Soc. London **223** A. 127 (1922).

[2] W. Noddack and I. Tacke, Naturwissenschaften **13,** 567 (1925).

[3] C. Perrier and E. Segré, J. Chem. Phys. **5,** 712 (1937); E. E. Motta, G. E. Boyd, and Q. V. Larson, Phys. Rev. **72,** 1270 L (1947).

[4] G. W. Parker, J. Reed, and J. W. Ruch, CNL–1, Clinton National Laboratory, Contract No. W–35–058–eng–71 (Jan. 9, 1948).

[5] C. Perrier and E. Segré, Nature **159,** 24 (1947).

[6] W. F. Meggers and B. F. Scribner, J. Research NBS **45,** 476 (1950) RP2161.

[7] H. Kayser, Tabelle der Schwingungszahlen, 1944 ed. (Edwards Brothers, Inc., Ann Arbor, Mich.)

[8] C. E. Moore, Atomic energy levels, NBS Circular 467, Vol. **II**, in press (1951).

[9] R. C. Gibbs and H. E. White, Proc. Nat. Acad. Sci. **14,** 559 (1928).

[10] H. Kayser and C. Runge, Wiedem. Ann. **52,** 93 (1894).

WASHINGTON, December 17, 1950.

Properties of Beryllium-Barium Titanate Dielectrics

Elmer N. Bunting, George R. Shelton, Ansel S. Creamer, and Bernard Jaffe

Dielectrics having compositions in the system BeO-$BaTiO_3$-TiO_2 were matured at 1,240° to 1,525° C. Data are given for the compositions, heat-treatments, absorption, and shrinkage. At 25° C the dielectric constant (K) and the reciprocal of the power factor (Q) were determined at frequencies of 50, 130, 1,000, and 20,000 kilocycles per second and at 3,000 megacycles per second when K is not greater than 50. At 1 megacycle per second K was determined at temperatures between −60° and +85° C and at 130 kilocycles per second, K and Q were determined at temperatures between 25° and 200° C. The linear thermal expansion was measured over the temperature range 25° to 700° C. It varied from 0.58 to 0.77 percent. The electrical resistivities of a few specimens were measured at 200° C, using a direct-current potential constantly applied for several days. In some cases the resistivity decreased by a factor of 10^3 over a few days. For specimens of some compositions, K and Q changed with time.

I. Introduction

Previous studies, by several investigators, of titanate dielectrics have shown that these ceramics are useful in the field of electronic instrumentation [1 to 5].[1] Special applications, such as in some miniature electronic devices, necessitate dielectrics in capacitors that will function at elevated temperatures without excessive electrical losses. One of the best insulators at high temperatures is beryllia, and some ceramics containing beryllia are characterized by low electrical losses [6]. Despite these facts and the increased interest in titanate ceramics during the past decade, very little information on beryllia-titania dielectrics has been published.

The present paper, dealing with dielectrics composed of beryllia, baria, and titania, is the fourth in a series on ceramic dielectrics made from titania and the oxides of the alkaline earth elements. An extended range in composition for the system BeO-BaO-TiO_2 was covered in this investigation in order to reveal dielectrics that may have desirable properties at various temperatures from −60° to +200° C.

II. Preparation of Dielectrics and Methods of Test

Dielectrics having the compositions shown in figure 1 were prepared from reagent quality barium carbonate and beryllia with the commercial grade of titania (grade TMO) used in the preparation of alkaline-earth titanate dielectrics previously investigated [2, 4, 5].

The methods of preparing these dielectrics and of determining their properties have been previously described [2, 4]. For measurements of K and Q at 25° to 200° C, the silvered test disk rested on a thin flat silver electrode on a hot-plate, and a silver-wire electrode, in a vertical position, touched the top center of the test disk. These silver electrodes were connected to the terminals on the Q-meter by short lengths (12 in.) of heavy copper wire. Temperatures of the disks on the hot-plate were measured with a calibrated copper-constantan thermocouple (B&S Gage No. 34) inserted into a small hole extending from the edge to the center of a similar disk placed next to the disk being measured. Due to the small thickness (0.1 in.) of the test specimen, the entire piece was maintained within a small temperature gradient. Tests made on specimens of barium titanate by this procedure gave a Curie point between 115° and 120° C, indicating an accuracy of about ±3° C.

For electrical leakage tests on capacitors, matured specimens 0.75 in. square by 0.010 in. thick were coated over an area of 2 cm^2 on each face with fired-on silver electrodes, and silver lead wires were attached. The specimens were heated in an electric furnace, and the temperature was measured on a mercury-in-glass thermometer placed next to the test piece. Leakage current was measured on a microammeter.

III. Results and Discussion

In table 1, data are given for the composition, heat-treatment, absorption, shrinkage, dielectric constant (K), and Q-value (reciprocal of the power factor) of mature specimens. The data for a given composition are considered to be the most representative of those obtained from measurements of four to seven specimens.

The effect of composition on the maturing temperature of the specimens is shown in figure 2, in which the specimens having the same maturing temperature are connected by lines. In all cases, the range in temperature within which mature specimens could be produced was not determined, but this range usually was not more than 25° C. Of the 33 bodies with ternary compositions, 25 were matured at 1,275° C, or less, within 1 or 2 hours. Specimens designated BBe63 (table 1) required several heat treatments of 4- to 6-hours duration at 1,315° C to reach maturity. One or more relatively low quin-

[1] Figures in brackets indicate the literature references at the end of this paper

tuple points in the phase diagram for this system are indicated by the low-maturing temperatures over a wide range in composition.

At 1 Mc/s and at 25° C, the variations in values of K and Q with composition are shown in figures 3 and 4, respectively. These diagrams were constructed from the data on K and Q given in table 1. In the binary system beryllia-titania, with the substitution of BeO for TiO_2 the values of K decreased from near 100 for TiO_2 to 17 for bodies of composition 6BeO:TiO_2. The observed value of K for BeO:TiO_2 was 57, compared to 71 reported by B. M. Wul [7]. The higher values sometimes found are probably due to the presence of small amounts of the lower oxide of titanium, which has a dielectric constant of several thousand. Q-values remained at several thousand until about 50 weight percent of BeO was present. Higher percentages of BeO resulted in specimens with lower Q, which decreased to 130 for 6BeO:TiO_2. The dielectric constant and Q-values for BeO (99.7% pure), matured at 1,925° C and measured at 25° C and 100 kc/s, have been reported as K=6.3 and Q=420 [6].

Within the ternary system, for dielectrics containing percentages of TiO_2 greater than about 60, the replacement of BaO by BeO did not change the values of K very much, but usually increased the Q-values. However, on the join 3BeO:TiO_2-BaO:2TiO_2, K decreased from 200 for BaO:2TiO_2 to 30 for 3BeO:TiO_2, with a minimum of 25 about midway along the join. The Q-values usually increased as the BeO content increased. On the 4BeO:TiO_2-2BaO:3TiO_2 join, K decreased rapidly from 900 for 2BaO:3TiO_2 to 23 for 4BeO:TiO_2. The addition of 9 weight percent of 4BeO:TiO_2 to 2BaO:3TiO_2 decreased Q from 50 for 2BaO:3TiO_2 to 33 for 2BBe49 (table 1), but greater percentages of 4BeO:TiO_2 increased the Q to a maximum of 3,000 at about 90 percent. With more than 90 percent, a decrease in Q to 300 for 4BeO:TiO_2 was found. On the 6BeO:TiO_2-$BaTiO_3$ join, K decreased from about 1,500 for $BaTiO_3$ to 17 for 6BeO:TiO_2. Q-values decreased from about 100 for $BaTiO_3$ to 30 when the content of 6BeO:TiO_2 was 25 percent (BBe67, table 1). Further additions of 6BeO:TiO_2, increased the Q to a maximum of 8,000 at 90 percent. However, at 100 percent of 6BeO:TiO_2, the Q had decreased to 130.

The effect of variation in frequency on the Q-values may also be noted in table 1. In the large majority of specimens, the Q-values were lower at 50 kc/s than at the higher frequencies. At 20 Mc/s, they were usually higher than at 1 Mc/s for specimens of high titania content, but lower for those of high beryllia or barium titanate content. At 3,000 Mc/s, higher values of Q were found than at 20 Mc/s for about half of the specimens tested.

In some miniaturized electronic equipment, the parts are subjected to temperatures of about 200° C, necessitating a knowledge of the properties of the dielectrics under these conditions. Data are given, in table 2, for K and Q measured at 130 kc/s at temperatures between 25° and 200° C, and for the average temperature coefficient of K. These data, in most instances, were obtained on one specimen only of each composition. For most of these dielectrics, K decreased with increasing temperature, and the average temperature coefficient of K ranged from −35 to −810 ppm/° C. Q-values usually were decreased greatly at higher temperatures, and at 200° C only seven specimens had Q greater than 500, with the highest (1,400) for 2BBe8. Only 10 of the 54 test pieces had a higher Q at 200° C than at 25° C, and most of the 10 were of high barium titanate content. The "Q" meter was warm and under voltage control of ±½ percent when the measurements were made on specimens having a capacitance of 100 to 200 $\mu\mu$f. Values of the average temperature coefficient of K are probably within ±20 ppm/° C. The temperature of the test pieces was raised at the rate of about 3 deg C a minute, and readings were taken at 25 deg C intervals. In order to illustrate the variations in Q resulting from changes in temperature and composition, figures 5 and 6 were constructed for temperatures of 100° and 200° C, respectively, and are based on the data in table 2.

In table 3 data are given for the values of K, at 1 Mc/s, over the temperature range of −60° to +85° C and for the average temperature coefficient of K. These values were obtained by measuring the capacitance of test disks at 10-deg intervals, with the temperature held constant at each interval for at least 15 min before measurements were made. The average values of the temperature coefficient of K, last column of table 3, are considered to be not better than ±10 ppm/° C, and most of them are negative within the range −40 to −840 ppm/° C. In order to illustrate the variation of K resulting from changes in temperature and composition, figures 7, 8, and 9 were constructed for temperatures of −60°, +100°, and +200° C, respectively. These figures contain isodielectric-constant lines, based on the data in tables 2 and 3.

After storage for 6 months under room conditions, two or more specimens of each composition were retested for K and Q at 25° C and 1 Mc/s. The results for the dielectrics that showed a significant change are given in table 4. These changes, particularly in K, are not as large, on the average, as those found to occur in the titanate systems previously studied [2, 4, 5]. Specimens of only one composition (BBe69), of high $BaTiO_3$ content, decreased in K and at the same time increased in Q by an appreciable amount. The absence of much solid solution between BeO and $BaTiO_3$ may account for the smaller change, with time, in the electrical properties of these dielectrics compared to those containing the other alkaline-earth oxides.

Linear thermal expansions, determined by the interferometer method, were moderately high for representative specimens, as shown in table 5.

In order to determine the usefulness of these dielectrics in capacitors at 200° C and under a constant potential, tests of the change in electrical resistivity over a period of time were made on some specimens of various compositions. On the assumption that practically all of the conductivity was

ugh the piece, the approximate volume resistivity was calculated from the amount of current and dimensions of the test piece between the fired-on electrodes. The data obtained are shown in table 6 and figures 10 and 11. Barium titanate specimens showed the fastest and largest decrease in resistivity, while the wafers of high beryllia content showed the least change in resistance. After the tests, only those specimens exhibiting a considerable decrease in resistivity were found to be colored black throughout the volume between the silver electrodes, while portions not under the electrodes remained unchanged in color. This decrease in resistivity and change in color is probably due to the reduction of the tetravalent titanium to the trivalent form. A reversal in the polarity of the applied voltage resulted in a temporary increase in resistivity, indicating a partial reoxidation. The most variable resistivity data were obtained for specimens of barium titanate. One of the factors most likely to influence the resistivity is the number of dark-colored spots in the specimen before the test. The data on barium titanate given in table 6 and figure 10 are the best values for specimens made with a commercial grade of titania. At 200° C and under 3-v direct current per mil, the resistivity for $BaTiO_3$ is 1.0×10^8 ohm-cm when measured immediately after applying the voltage. For the other specimens given in table 6, the resistivity just after application of the voltage is greater than 10^{10} ohm-cm. In order to find how much voltage was needed to cause a decrease in resistivity with time at 200° C, the voltage on specimen BBe67 was increased to 100 (10 v/mil) after the specimen had been at 30 v, and 200° C, for 21 days. The resistivity remained greater than 10^{10} ohm-cm for 4 additional days and then decreased to 2×10^7 ohm-cm.

The decrease in resistivity shown by these dielectrics, when subjected to a constant voltage, greatly lowers their usefulness for high-temperature duty. Those bodies, however, which showed a small loss in resistivity at 200° C could be used for the production of capacitors for operation at this temperature if the working voltage is not too high.

IV. Summary

Dielectrics having compositions in the system $BeO\text{-}BaTiO_3\text{-}TiO_2$ can be prepared from mixtures of titania with barium carbonate and beryllia. Mature dielectrics (less than 0.1% absorption) were made by heating specimens to various temperatures within the range 1,240° to 1,525° C.

The dielectric constant (K) varied from 16 for specimens with compositions near that of $6BeO\text{:}TiO_2$, to several hundred for dielectrics of high baria content. Most of the temperature coefficients of K were negative. The Q-values, at 25° C, ranged from 15 to 10,000 and were low for dielectrics containing large percentages of baria and for those with high beryllia content in the binary system $BeO\text{-}TiO_2$. At higher temperatures, most of the Q-values decreased greatly. K and Q-values of some of the dielectrics changed with time.

Values for the linear thermal expansion over the temperature range 25° to 700° C, varied from 0.58 to 0.77 percent and were obtained on dielectrics of widely different compositions.

When subjected to an electrical potential at 200° C for many hours, some of the dielectrics decreased greatly in electrical resistivity.

V. References

[1] A. von Hippel et al., Ind. Eng. Chem. **38**, 1097 (1946).
[2] E. N. Bunting, G. R. Shelton, and A. S. Creamer, J. Research NBS **38**, 337 (1947) RP1776; J. Am. Ceram. Soc. **30**, 4 (1947).
[3] U. S. Pat. 2,420,692, by E. Wainer.
[4] G. R. Shelton, A. S. Creamer, and E. N. Bunting, J. Research NBS **41**, 17 (1948) RP1899.
[5] E. N. Bunting, G. R. Shelton, and A. S. Creamer, J. Research NBS **43**, 237 (1949) RP2025.
[6] R. F. Geller, P. J. Yavorsky, B. L. Steirman, and A. S. Creamer, J. Research NBS **36**, 277 (1946) RP1703.
[7] B. M. Wul, Nature **156**, 480 (1945).

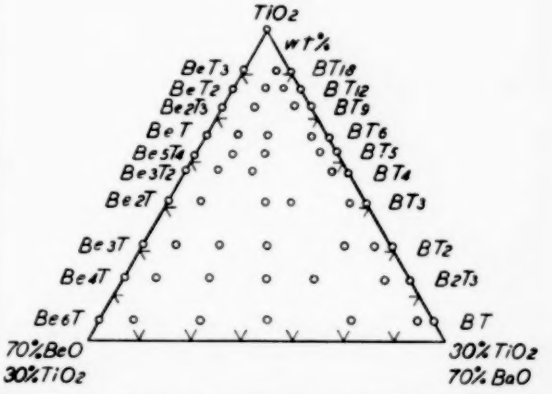

FIGURE 1. *Ternary diagram for system* $BeO\text{-}BaTiO_3\text{-}TiO_2$ *showing compositions studied.*

B=BaO; Be=BeO; T=TiO_2; thus Be6T=6BeO:TiO_2.

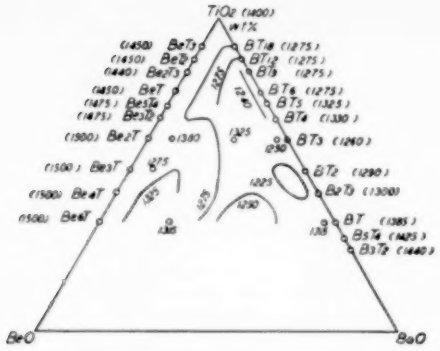

FIGURE 2. *Approximate maturing temperature (°C) after calcining treatment.*

B = BaO; Be = BeO; T = TiO_2.

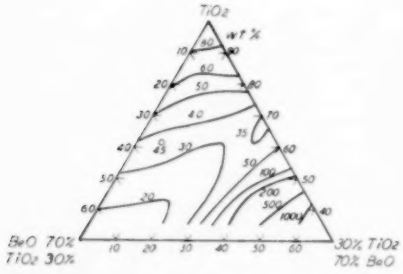

FIGURE 3. *Dielectric constant, at 25° C and 1 Mc/s, with varying composition within the ternary system* BeO-$BaTiO_3$-TiO_2.

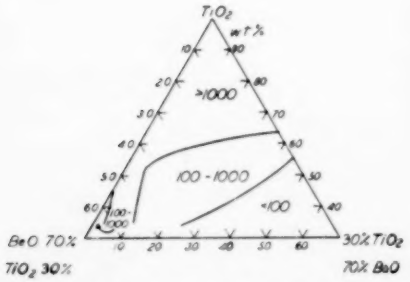

FIGURE 4. *Q-values, at 25° C and 1 Mc/s, with varying composition within the ternary system* BeO-$BaTiO_3$-TiO_2.

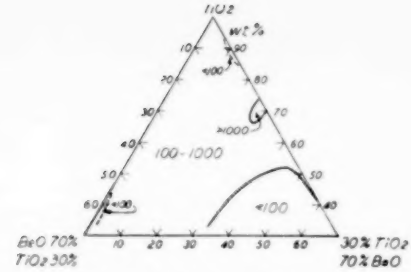

FIGURE 5. *Q-values, at 100° C and 130 kc/s, with varying composition within the ternary system* BeO-$BaTiO_3$-TiO_2.

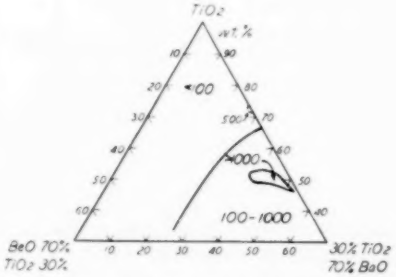

FIGURE 6. *Q-values, at 200° C and 130 kc/s, with varying composition within the ternary system* BeO-$BaTiO_3$-TiO_2.

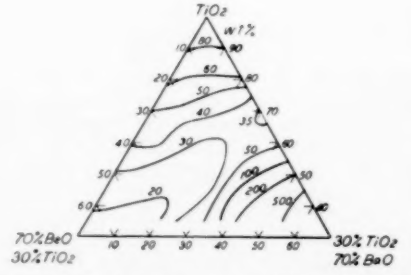

FIGURE 7. *Dielectric constant, at −60° C and 1 Mc/s, with varying composition within the ternary system* BeO-$BaTiO_3$-TiO_2.

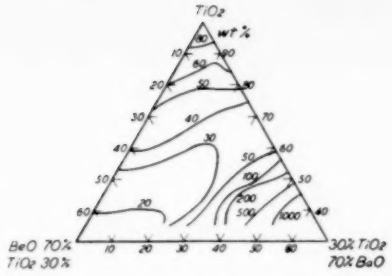

FIGURE 8. *Dielectric constant, at 100° C and 130 kc/s, with varying composition within the ternary system* $BeO\text{-}BaTiO_3\text{-}TiO_2$.

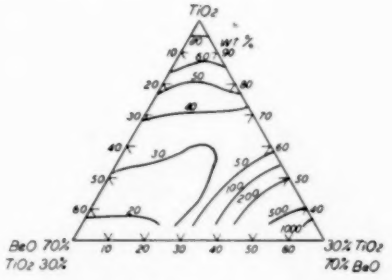

FIGURE 9. *Dielectric constant, at 200° C and 130 kc/s, with varying composition within the ternary system* $BeO\text{-}BaTiO_3\text{-}TiO_2$.

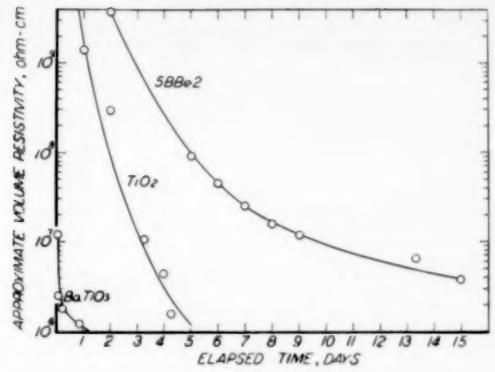

FIGURE 10. *Approximate volume resistivity for specimens of* $BaTiO_3$, TiO_2, *and* $5BBe_2$, *at 200° C and 30 volts direct-current, after varying lengths of time.*

Thickness of all specimens was approximately 0.010 in.

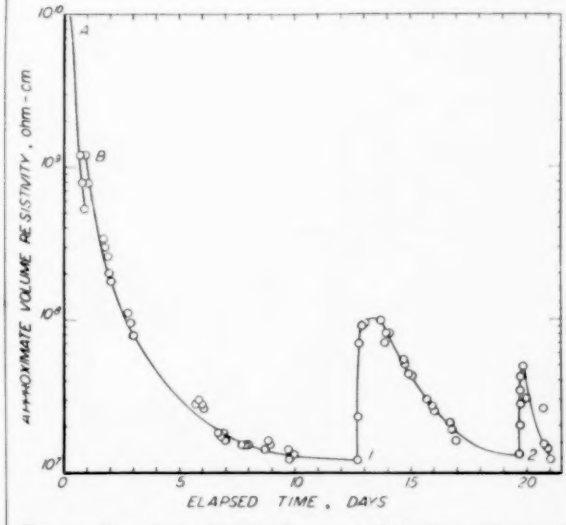

FIGURE 11. *Approximate volume resistivity for specimens of* $5BBe_2$, *0.010 in. thick, at 200° C and 110 or 30 volts direct-current, after varying lengths of time.*

A, 110 v; B, 30 v; 1, polarity reversed; 2, original polarity restored.

TABLE 1. *Composition, heat-treatment, absorption, shrinkage, dielectric constant K, and Q of bodies in the system* $6BeO:TiO_2$-$BaTiO_3$-TiO_2

Specimen designation	Proportion of end members of join, weight percent		Composition weight			Heat treatment				Absorption	Shrinkage	Dielectric constant, K, at 25° C				Reciprocal, Q, of power factor at 25° C			
						No. 1		No. 2											
			BeO	BaO	TiO_2	Temperature	Time	Temperature	Time			50 kc/s	1 Mc/s	20 Mc/s	3,000 Mc/s	50 kc/s	1 Mc/s	20 Mc/s	3,000 Mc/s
	$BeO:3TiO_2$	$BaO:18TiO_2$																	
			%	%	%	°C	hr	°C	hr	%	%								
BeT3	100.0	0.0	9.45		90.55	1,450	1			0.00	20	80	80	80		2,000	10,000	10,000	
18BBe6	31.75	68.25	3.0	6.55	90.45	1,275	1			.00	22	77	77	77		800	5,000	10,000	
BT18	0.0	100.0		9.6	90.4	1,100	1	1,275	1	.00	15	75	74	74		600	2,000	7,000	
	$BeO:2TiO_2$	$BaO:12TiO_2$																	
BeT2	100.0	0.0	13.5		86.5	1,450	1			.00	21	71	71	71		1,700	10,000	10,000	
12BBe4	52.9	47.1	7.2	6.5	86.3	1,275	1			.02	20	65	65	65		800	4,000	5,000	
12BBe7	22.2	77.8	3.0	10.7	86.3	1,100	1	1,240	1	.07	13	64	64	64		260	1,200	2,000	
BT12	0.0	100.0		13.8	86.2	1,100	1	1,275	1	.00	15	68	68	68		460	1,800	2,000	
	$2BeO:3TiO_2$	$BaO:9TiO_2$																	
Be2T3	100.0	0.0	17.3		82.7	1,450	1			.00	22	66	66	66		7,000	10,000	10,000	
9BBe3	63.0	37.0	10.9	6.6	82.5	1,275	1			.00	20	57	57	57		1,000	3,500	3,500	
9BBe8	17.4	82.6	3.0	14.5	82.5	1,240	1			.07	21	60	60	60		1,000	4,000	4,000	
BT9	0.0	100.0		17.6	82.4	1,100	1	1,275	1	.00	15	60	60	60		1,000	2,000	5,000	
	$BeO:TiO_2$	$BaO:6TiO_2$																	
BeT	100.0	0.0	23.85		76.15	1,450	1			.00	20	57	57	57		1,000	8,000	10,000	
6BBe2	73.2	26.8	17.45	6.5	76.05	1,275	1			.00	19	51	51	51		1,000	3,000	3,000	
6BBe5	50.0	50.0	11.9	12.1	76.0	1,275	1			.00	15	49	49	49		1,000	3,000	3,000	
6BBe8	12.6	87.4	3.0	21.2	75.8	1,100	1	1,240	1	.00	17	48	48	48	46	2,000	3,000	3,000	3,800
BT6	0.0	100.0		24.3	75.7	1,100	1	1,275	1	.01	14	46	45	44	44	1,100	3,000	2,200	1,200
	$5BeO:4TiO_2$	$BaO:5TiO_2$																	
Be5T4	100.0	0.0	28.1		71.9	1,475	1			.00	22	52	52	52		2,000	10,000	10,000	
5BBe2	76.6	23.4	21.5	6.5	72.0	1,275	1			.00	20	48	48	48	48	1,000	4,000	4,000	360
5BBe5	50.0	50.0	14.1	13.9	72.1	1,100	1	1,275	1	.02	17	43	43	43	42	500	1,700	1,700	1,600
5BBe8	10.7	89.3	3.0	24.8	72.2	1,100	1	1,240	1	.03	17	42	42	42	41	2,000	4,000	7,000	2,600
BT5	0.0	100.0		27.7	72.3	1,100	1	1,325	1	.04	12	39	37	37	39	1,500	2,000	2,000	2,100
	$3BeO:2TiO_2$	$BaO:4TiO_2$																	
Be3T2	100.0	0.0	32.0		68.0	1,475	1			.01	22	45	45	45	44	1,600	10,000	10,000	1,500
4BBe2	80.0	20.0	25.55	6.5	67.95	1,275	1			.00	21	41	41	41	44	800	2,500	5,000	1,600
4BBe4	57.8	42.2	18.7	13.7	67.6	1,275	1			.00	16	41	41	41	39	800	4,000	4,500	6,000
4BBe9	9.4	90.6	3.0	29.4	67.6	1,240	1			.00	22	36	36	36	35	2,000	1,600	1,000	6,000
BT4	0.0	100.0		32.4	67.6	1,100	1	1,330	1	.03	17	34	34	35	33	2,000	2,000	3,700	2,600
	$2BeO:TiO_2$	$BaO:3TiO_2$																	
Be2T	100.0	0.0	38.5		61.5	1,500	1			.00	22	40	40	40		2,000	7,000	7,000	
3BBe1	83.3	16.7	32.1	6.5	61.4	1,100	1	1,380	2	.01	16	44	44	44		1,000	4,000	6,000	
3BBe3	50.0	50.0	19.3	19.5	61.2	1,100	1	1,275	1	.03	16	31	31	31	30	220	1,600	2,000	3,000
3BBe5	39.0	61.0	15.0	23.8	61.2	1,100	1	1,325	1	.01	16	30	30	30		700	1,000	1,000	
3BBe6	7.8	92.2	3.0	35.9	61.1	1,100	1	1,250	1	.00	17	37	37	37	38	1,000	1,000	1,000	2,500
3BBe9	0.0	100.0		39.0	61.0	1,100	1	1,260	1	.03	17	44	44	43	42	720	650	800	460
	$3BeO:TiO_2$	$BaO:2TiO_2$																	
Be3T	100.0	0.0	48.4		51.6	1,500	1			.01	20	30	30	30		700	6,000	6,000	
2BBe1	86.7	13.3	42.0	6.5	51.5	1,275	1	1,275	1	.00	19	29	29	29	29	200	800	2,000	1,700
2BBe3	70.0	30.0	33.9	14.7	51.4	1,100	1	1,325	2	.05	18	25	25	25		200	330	470	
2BBe5	50.0	50.0	24.2	24.5	51.3	1,100	1	1,275	1	.00	18	23	23	23	21	230	1,000	1,000	1,800
2BBe8	18.6	81.4	9.0	39.9	51.1	1,100	1	1,225	1	.00	16	70	70	70		200	130	80	
2BBe9	6.2	93.8	3.0	45.9	51.1	1,100	1	1,225	1	.00	16	155	155	155		200	62	35	
BT2	0.0	100.0		49.0	51.0	1,100	1	1,260	1	.02	16	204	200	197		100	70	35	

TABLE 1. *Composition, heat-treatment, absorption, shrinkage, dielectric constant K, and Q of bodies in the system* $6BeO:TiO_2$-$BaTiO_3$-TiO_2—Continued.

	Proportion of end members of join, weight percent		Composition weight			Heat treatment				Absorption	Shrinkage	Dielectric constant, K, at 25° C				Reciprocal, Q, of power factor at 25° C			
						No. 1		No. 2											
	$4BeO:TiO_2$	$2BaO:3TiO_2$	BeO	BaO	TiO_2	Temperature	Time	Temperature	Time			50 kc/s	1 Mc/s	20 Mc/s	3,000 Mc/s	50 kc/s	1 Mc/s	20 Mc/s	3,000 Mc/s
			%	%	%	°C	hr	°C	hr	%	%								
Be4T	100.0	0.0	55.6		44.4	1,500	1			.00	21	23	23	23		90	300	440	
1Be41	88.4	11.6	49.2	6.5	44.3	1,100	1	1,325	1	.00	18	23	23	23	23	900	3,000	3,000	2,000
2Be43	70.0	30.0	38.9	16.8	44.2	1,100	1	1,325	3	.01	16	21	21	21		300	600	1,000	
3Be45	50.0	50.0	27.8	28.1	44.1	1,100	1	1,275	1	.01	16	27	27	27	24	300	270	190	73
4Be47	25.0	75.0	13.9	42.1	44.0	1,100	1	1,250	2	.01	16	170	165	155		53	36	20	
5Be49	9.0	91.0	5.0	51.1	43.9	1,100	1	1,225	1	.00	16	490	470	450		45	33	18	
B2T3	0.0	100.0		56.1	43.9	1,250	1	1,300	1	.01	11	910	900	890		70	50	25	
	$6BeO:TiO_2$	$BaO:TiO_2$																	
Be6T	100.0	0.0	65.2		34.8	1,525	1			.07	21	17	17	17		40	130	110	
1Be61	90.1	9.9	58.8	6.5	34.7	1,100	1	1,325	3	.08	18	16	16	16		800	8,000	10,000	
3Be63	70.0	30.0	45.7	19.7	34.6	1,100	1	1,315	14	.08	15	17	17	17	15	110	130	600	950
5Be65	50.0	50.0	32.6	32.9	34.5	1,100	1	1,250	1	.00	18	65	64	61		91	68	40	
7Be67	25.0	75.0	16.3	49.3	34.4	1,100	1	1,250	2	.02	14	680	660	620		49	30	15	
9Be69	4.6	95.4	3.0	62.7	34.3	1,100	1	1,325	2	.00	14	1,170	1,140	1,120		86	60	26	
BT	0.0	100.0		65.7	34.3	1,245	1	1,385	2	.05	11	1,500	1,500	1,500		80	100	50	

TABLE 2. *Dielectric constant (K), Q-values, and average temperature coefficient of K from 25° to 200°C, at 130 kc/s*

Specimen designation	K at °C—						Q at °C—						Average temperature coefficient of K
	25°	100°	125°	150°	175°	200°	25°	100°	125°	150°	175°	200°	
													ppm/°C
TiO_2	98	92	90	89	88	86	3,000	670	290	140	112	43	−770
BeT3	78.2	74.8	73.4	72.4	71.1	70.2	4,000	670	420	230	91	48	−700
18BBe6	77.8	73.9	73.1	72.2	71.9	.....	1,200	190	100	46	20	.....	−525
BT18	73.2	71.7	71.0	70.6	69.2	.....	560	46	26	14	8	.....	−550
BeT2	68.8	64.7	63.5	62.4	61.6	60.7	4,000	630	370	160	84	38	−720
12BBe4	65.9	62.6	61.7	61.0	.....	.....	1,400	160	69	30	.....	.....	−620
12BBe7	59.1	56.7	56.0	55.6	55.3	.....	300	240	100	52	23	.....	−440
BT12	69.6	66.8	66.3	66.0	.....	.....	850	59	31	19	.....	.....	−420
Be2T3	66.4	62.5	61.3	60.4	59.8	.....	4,000	340	120	46	21	.....	−700
9BBe3	57.8	54.8	54.0	53.5	.....	.....	1,100	200	80	34	.....	.....	−610
9BBe8	61.2	58.8	58.2	57.7	.....	.....	460	140	57	25	.....	.....	−460
BT9	57.4	55.4	54.8	54.4	54.2	54.2	2,000	210	89	48	27	16	−340
BeT	54.2	51.3	50.6	49.9	49.4	.....	2,000	260	125	45	16	.....	−610
6BBe2	50.8	48.6	48.1	47.6	47.4	.....	1,500	180	81	40	19	.....	−470
6BBe5	49.5	47.7	47.2	46.9	46.5	46.4	240	250	180	100	43	21	−370
6BBe8	48.5	47.3	47.0	46.7	46.6	.....	1,400	270	130	55	22	.....	−260
BT6	48.8	48.0	47.9	47.8	47.8	.....	1,000	71	29	17	9	.....	−140
Be5T4	53.4	50.5	49.9	49.2	48.5	48.1	5,000	820	520	290	130	63	−590
5BBe2	48.3	45.8	45.1	44.7	44.2	43.9	2,600	320	160	80	40	20	−540
5BBe5	43.3	42.0	41.6	41.3	41.2	41.0	710	280	140	63	27	15	−310
5BBe8	42.1	41.4	41.2	41.2	41.1	41.1	430	280	170	95	46	24	−130
BT5	35.7	35.4	35.2	35.0	35.0	35.0	2,000	1,400	1,400	1,100	800	500	−120
Be3T2	43.8	41.6	41.0	40.6	40.3	.....	3,000	350	170	69	34	.....	−810
4BBe2	41.0	39.1	38.7	38.3	38.1	.....	900	230	93	42	18	.....	−490
4BBe4	41.3	40.0	39.6	39.4	39.3	39.3	2,000	320	130	62	28	14	−290
4BBe9	36.9	36.2	36.0	36.1	36.2	36.5	300	400	300	200	72	35	.....
BT4	35.6	35.5	35.5	35.5	35.6	35.7	3,000	720	330	155	78	35	.....
Be2T	44.5	41.2	40.6	40.4	39.8	39.6	10,000	280	125	64	27	13	−610
3BBe1	46.3	43.8	43.2	42.5	42.4	42.2	800	330	130	54	24	12	−550
3BBe5	30.9	30.4	30.3	30.2	30.2	30.2	620	520	320	180	87	45	−130
3BBe6	31.0	30.9	30.9	30.8	30.8	30.8	280	330	290	210	140	83	− 35
3BBe9	38.1	37.1	37.4	37.0	36.8	36.6	300	280	290	250	210	160	−230
BT3	44.5	43.8	43.6	43.5	43.2	42.9	700	600	600	600	600	600	−210
Be3T	32.5	31.3	31.0	30.6	30.6	30.8	4,000	300	130	50	22	12	.....
2BBe1	29.1	28.0	27.6	27.4	27.3	27.2	170	410	210	91	41	20	−380
2BBe3	26.0	26.0	26.0	26.1	26.1	26.15	300	110	86	59	35	22	+50
2BBe5	22.6	22.4	22.4	22.4	22.4	22.5	170	350	300	210	110	59	.....
2BBe8	70.4	73.8	76.5	75.9	71.2	66.4	160	210	390	800	1,500	1,400	.....
2BBe9	163	188	212	195	160	136	62	88	170	470	1,000	1,000	.....
BT2	183	222	222	174	145	125	130	220	320	440	500	600	(a)
Be4T	22.9	24.2	24.6	24.8	24.9	24.9	51	24	29	31	27	18	+480
2BBe41	23.6	23.1	23.0	23.0	22.8	22.8	3,000	490	210	93	35	22	−260
2BBe43	20.5	20.5	20.5	20.5	20.6	20.6	83	91	96	93	79	60	+30
2BBe45	27.2	27.3	27.6	27.6	27.4	27.2	140	130	150	140	130	110	.....
2BBe47	161	199	264	215	163	97	44	56	110	200	190	110	.....
2BBe49	482	586	656	550	379	286	30	59	132	320	530	640	.....
B2T3	796	1,130	961	661	487	403	60	170	360	550	700	900	(a)
Be6T	16.8	17.0	17.1	17.2	17.3	17.4	68	49	49	44	40	19	+210
BBe61	16.4	16.2	16.1	16.0	16.0	16.1	150	190	160	120	68	38	.....
BBe63	17.3	17.3	17.3	17.2	17.3	17.3	120	110	100	99	97	81	.....
BBe65	64.0	71.6	76.1	72.3	62.5	56.0	78	100	160	240	210	150	.....
BBe67	680	906	1,190	900	580	428	33	56	150	380	680	730	.....
BBe69	1,190	1,320	4,140	2,100	1,250	940	75	63	67	92	67	52	.....
BT	1,210	1,850	2,870	1,550	1,100	820	110	47	1,000	1,000	590	190	(a)

(a) Maximum K value occurs between 100° and 125°C.

TABLE 3. *Dielectric constant, at 1 Mc/s, from −60° to +85°C, and average temperature coefficient of dielectric constant*

Specimen designation	Values of K at °C—															Average temperature coefficient of K
	−60°	−50°	−40°	−30°	−20°	−10°	0°	+10°	20°	30°	40°	50°	60°	70°	80°	85°	
																	ppm/°C
BeT3	84.8	83.9	82.9	82.0	81.2	80.4	79.8	79.2	78.8	78.4	77.7	77.0	76.4	75.8	75.3	75.1	−840
3BBe6	82.4	81.6	80.8	80.1	79.4	78.8	78.2	77.6	77.1	76.6	75.9	75.4	74.8	74.2	73.7	73.4	−790
BT18	80.5	79.7	78.9	78.3	77.6	77.0	76.5	76.0	75.4	74.8	74.3	73.8	73.4	73.0	72.6	72.5	−720
BeT2	75.9	75.0	74.2	73.4	72.7	72.1	71.5	71.0	70.4	69.9	69.4	68.8	68.2	67.7	67.2	67.0	−850
2BBe4	70.2	69.5	68.8	68.1	67.5	67.0	66.5	66.0	65.6	65.2	64.7	64.2	63.8	63.3	62.9	62.7	−770
2BBe7	70.4	69.8	69.1	68.7	68.1	67.7	67.2	66.8	66.5	66.1	65.7	65.3	64.8	64.5	64.1	63.9	−670
BT12	73.2	72.5	71.9	71.3	70.8	70.3	69.9	69.5	69.1	68.8	68.4	68.0	67.6	67.3	67.0	66.8	−620
Be2T3	71.2	70.5	69.7	69.0	68.4	67.8	67.2	66.7	66.1	65.6	65.1	64.6	64.1	63.7	63.2	63.0	−830
3BBe3	61.2	60.6	60.1	59.6	59.0	58.6	58.1	57.7	57.3	56.9	56.4	56.0	55.6	55.2	54.8	54.6	−770
9BBe8	64.2	63.6	63.1	62.7	62.4	62.0	61.7	61.3	61.0	60.6	60.3	60.0	59.7	59.4	59.1	58.9	−600
BT9	63.2	62.8	62.4	62.0	61.7	61.3	61.0	60.7	60.4	60.0	59.8	59.5	59.2	59.0	58.7	58.6	−480
BeT	57.3	56.7	56.2	55.7	55.2	54.7	54.3	53.9	53.4	53.0	52.6	52.2	51.8	51.5	51.1	50.9	−800
6BBe2	53.7	53.3	52.9	52.5	52.1	51.7	51.4	51.0	50.6	50.3	49.9	49.6	49.3	49.0	48.7	48.6	−700
6BBe5	51.1	50.8	50.4	50.1	49.8	49.5	49.2	49.0	48.8	48.5	48.3	48.0	47.8	47.6	47.4	47.3	−520
6BBe8	49.2	49.0	48.8	48.6	48.4	48.2	48.1	47.9	47.8	47.6	47.5	47.3	47.1	47.0	46.9	46.7	−360
BT6	46.5	46.3	46.2	46.0	45.8	45.6	45.4	45.3	45.1	45.0	44.9	44.8	44.7	44.6	44.5	44.5	−300
Be5T4	56.2	55.5	54.9	54.4	53.9	53.5	53.1	52.8	52.4	52.0	51.6	51.2	50.9	50.5	50.1	50.0	−800
5BBe2	50.8	50.4	50.0	49.7	49.3	49.0	48.6	48.3	48.0	47.6	47.3	47.0	46.7	46.4	46.1	46.0	−680
5BBe5	44.6	44.3	44.0	43.8	43.6	43.4	43.2	43.0	42.8	42.6	42.4	42.2	42.0	41.9	41.7	41.6	−870
5BBe8	42.4	42.2	42.1	42.0	41.9	41.8	41.7	41.6	41.6	41.5	41.4	41.3	41.3	41.2	41.1	41.1	−220
BT5	36.7	36.7	36.7	36.7	36.7	36.7	36.7	36.7	36.7	36.7	36.7	36.7	36.7	36.7	36.7	36.7	0
Be3T2	46.3	45.9	45.4	45.0	44.6	44.2	43.9	43.7	43.4	43.2	42.9	42.6	42.3	42.0	41.7	41.6	−730
4BBe2	42.9	42.5	42.1	41.8	41.5	41.2	41.0	40.7	40.5	40.2	40.0	39.7	39.5	39.2	39.0	38.9	−670
4BBe4	42.4	42.2	41.9	41.7	41.5	41.3	41.1	41.0	40.8	40.6	40.4	40.3	40.1	39.9	39.8	39.7	−440
4BBe9	35.5	35.4	35.4	35.4	35.4	35.3	35.3	35.3	35.3	35.3	35.3	35.3	35.3	35.3	35.2	35.2	−50
BT4	33.0	33.1	33.2	33.3	33.4	33.5	33.6	33.7	33.7	33.8	33.9	33.9	34.0	34.0	34.0	34.0	+200
Be2T	43.4	43.0	42.6	42.2	41.9	41.6	41.3	41.1	40.8	40.5	40.2	40.0	39.7	39.4	39.2	39.1	−720
3BBe1	46.9	46.5	46.0	45.5	45.2	44.8	44.4	44.2	43.9	43.6	43.3	43.0	42.7	42.4	42.2	42.0	−750
3BBe5	31.1	31.0	30.9	30.8	30.7	30.7	30.6	30.6	30.5	30.4	30.4	30.3	30.3	30.2	30.2	30.2	−280
3BBe6	30.0	30.0	30.0	30.0	30.0	30.0	29.9	29.9	29.9	29.9	29.9	29.9	29.9	29.9	29.9	29.8	−40
3BBe9	36.8	36.8	36.8	36.8	36.7	36.7	36.7	36.7	36.7	36.7	36.7	36.7	36.6	36.6	36.6	36.6	−40
BT3	44.6	44.5	44.5	44.4	44.3	44.3	44.2	44.1	44.1	44.0	43.9	43.9	43.8	43.8	43.7	43.7	−140
Be3T	30.5	30.2	30.0	29.8	29.6	29.4	29.3	29.1	29.0	28.9	28.7	28.6	28.5	28.4	28.3	28.3	−510
2BBe1	29.9	29.7	29.5	29.3	29.1	28.9	28.8	28.7	28.5	28.4	28.3	28.1	28.0	27.9	27.7	27.7	−510
2BBe3	24.6	24.6	24.5	24.5	24.5	24.4	24.4	24.4	24.4	24.3	24.3	24.3	24.3	24.2	24.2	24.2	−120
2BBe5	22.5	22.5	22.5	22.5	22.6	22.6	22.6	22.6	22.6	22.6	22.6	22.6	22.6	22.6	22.6	22.6	+20
2BBe8	63.2	64.0	64.8	65.5	66.2	66.8	67.5	68.1	68.7	69.2	69.6	69.9	70.1	70.4	70.9	71.3	+820
2BBe9	126	128	134	138	142	146	150	153	157	159	161	163	165	167	171	173	+2,180
BT2	163	168	173	177	182	187	192	196	200	202	204	207	210	214	219	221	+2,080
Be4T	24.5	24.3	24.2	24.0	23.9	23.8	23.7	23.6	23.6	23.5	23.4	23.4	23.3	23.3	23.4	23.4	−310
2BBe41	23.8	23.7	23.6	23.5	23.4	23.4	23.3	23.2	23.2	23.1	23.0	23.0	22.9	22.9	22.8	22.8	+50
2BBe43	20.5	20.5	20.5	20.5	20.5	20.5	20.5	20.5	20.5	20.6	20.6	20.6	20.6	20.6	20.6	20.6	+220
2BBe45	26.0	26.1	26.2	26.3	26.3	26.4	26.5	26.5	26.6	26.6	26.7	26.7	26.8	26.8	26.8	26.9	+220
2BBe47	120	125	132	137	143	148	154	159	164	167	170	173	177	182	187	191	+3,150
2BBe49	311	330	350	365	385	405	430	450	470	480	490	500	510	520	535	543	+3,750
B2T3	600	630	660	690	720	745	780	825	870	885	900	910	920	935	960	980	+3,300
Be6T	16.5	16.5	16.4	16.4	16.4	16.3	16.3	16.3	16.2	16.2	16.2	16.2	16.2	16.2	16.2	16.2	−120
BBe61	16.5	16.5	16.5	16.4	16.4	16.4	16.3	16.3	16.3	16.3	16.2	16.2	16.2	16.2	16.2	16.2	−160
BBe63	16.7	16.7	16.7	16.7	16.7	16.8	16.8	16.8	16.8	16.8	16.8	16.8	16.8	16.8	16.8	16.8	+40
BBe65	53.8	55.1	56.5	57.8	59.0	60.1	61.2	62.1	63.2	64.1	64.8	65.4	66.0	66.6	67.4	67.9	+1,600
BBe67	419	450	480	510	540	570	600	630	660	680	705	710	730	744	775	790	+4,260
BBe69	638	660	680	710	740	780	830	910	1,140	1,105	1,055	1,015	1,005	1,000	1,030	1,048	.....
BT	805	815	860	910	980	1,060	1,160	1,310	1,570	1,420	1,360	1,330	1,300	1,330	1,400	1,440	.....

Specimen designation	Dielectric constant, K				Quality factor, Q			
	After 1 day	After 6 months	Change	Average change	After 1 day	After 6 months	Change	Average change
Specimens with significant change in K								
			Percent	*Percent*			*Percent*	*Percent*
2BBe9-2	161	157	−2		49	51	+4	
-3	156	154	−1	−2	55	53	−4	0
-4	152	150	−1		62	62	0	
2BBe47	167	165	−1		32	19	−40	
-1	190	186	−2	−2	28	29	+4	−10?
-B	177	173	−2		36	38	+6	
2BBe49	479	463	−3		30	30	0	
-3	462	450	−3	−3	31	31	0	0
-4	454	444	−2		33	33	0	
BBe67	692	654	−4		27	24	−11	
-2	695	645	−3	−2½	29	29	0	−4?
-3	671	668	−½		27	24	−11	
BBe69-2	1,110	1,080	−3		39	40	+2	
-4	1,120	1,096	−2	−3½	58	64	+10	
-5	1,097	1,060	−3		60	66	+10	+7
-8	1,140	1,074	−6		53	55	+4	
Specimens with significant change in Q								
3BBe5	30.5	30.6			830	180	−78	
-1	30.5	30.1			1,600	600	−63	−70
3BBe6-1	28.4	[a] 29.5			1,000	31	−97	
-2	28.9	29.2			430	55	−87	−92
3BBe9	37.4	[a] 36.8			1,100	470	−57	
-A	36.7	36.9			200	91	−55	−59
-B	36.9	37.2			140	51	−64	
2BBe3	24.6	[a] 27.0			330	8	−98	
-B	24.4	24.4			160	92	−43	−70
2BBe5-1	25.0	25.3			1,000	800	−20	
-A	22.0	22.0			400	260	−35	−24
-C	22.6	22.7			400	330	−18	
BBe43-1	18.6	[a] 19.3			600	28	−95	
-A	20.1	20.3			150	26	−83	−90
BBe45	26.0	26.0			270	100	−30	
-A	26.5	26.7			180	62	−66	−48
BBe65	63.0	63.2			68	26	−62	
-A	62.6	62.4			58	52	−10	−26
-B	62.6	62.3			59	55	−7	

[a] Apparent change not significant due to small capacitance of specimen.

TABLE 4.

Changes in K and Q of specimens after 6 months storage

(Measured at 1 Mc/s and 25° C)

Specimen designation	Temperature range from 25° C to—						
	100°	200°	300°	400°	500°	600°	700°
	Percent	*Percent*	*Percent*	*Percent*	*Percent*	*Percent*	*Percent*
$6BeO:TiO_2$	0.04	0.11	0.19	0.28	0.38	0.48	0.58
$4BeO:TiO_2$	.04	.12	.20	.33	.46	.58	.69
$2BeO:TiO_2$	.05	.12	.21	.30	.39	.49	.59
$BeO:TiO_2$	.04	.11	.21	.33	.49	.68	.77
5BBe5	.05	.14	.24	.33	.43	.53	.64
BBe65	.05	.13	.22	.33	.43	.54	.65

TABLE 5.

Linear thermal expansion

Specimen designation	Resistivity in ohm-cm after—				
	1 hr	1 day	2 days	5 days	21 days
$BaTiO_3$	6×10^{9}	8×10^{9}	[a] 8×10^{9}		
TiO_2	$>10^{10}$	10^{9}	2×10^{9}	10^{9}	[b] 6×10^{9}
5BBe2	$>10^{10}$	$>10^{10}$	4×10^{9}	10^{9}	[c] 5×10^{9}
2BBe45	$>10^{10}$	$>10^{10}$	$>10^{10}$	$>10^{10}$	$>10^{10}$
BBe61	$>10^{10}$	$>10^{10}$	$>10^{10}$	$>10^{10}$	$>10^{10}$
BBe67	$>10^{10}$	$>10^{10}$	$>10^{10}$	$>10^{10}$	$>10^{10}$

TABLE 6. *Approximate volume resistivity of some specimens, 0.010 in. thick, measured at 200° C and at a potential of 30-v direct current maintained for various periods of time*

[a] In 2 hr after 2 days, the resistivity began to decrease rapidly and, within an additional 2 hr had decreased to 1×10^{8} ohm-cm, when test was discontinued.

[b] After 7 days.

[c] After 15 days.

WASHINGTON, December 19, 1950.

of Research of the National Bureau of Standards Vol. 47, No. 1, July 1951 Research Paper 2223

Refractivity of Lithium Fluoride With Application to the Calibration of Infrared Spectrometers[1]

Leroy W. Tilton and Earle K. Plyler

The refractivity of lithium fluoride was measured from 0.4 to 5.9 microns and expressed for 23.6° C by a four-constant dispersion formula by which an interpolation table was prepared for included wavelengths. The temperature coefficient was determined as -1.63×10^{-5} (constant within ±0.05) for the visible region and a table of indices made for 20° to 60° C at equal intervals of λ from 0.4 to 0.7 micron. Most of the observed indices agree with the tabulated values within $\pm3\times10^{-6}$ for the visible region and $\pm3\times10^{-5}$ in the infrared. In calibrating an infrared spectrometer simple empirical relationships between drum reading, D, and wavelength are not entirely reliable. Three steps are recommended: (1) the expression of D in terms of E, the emergent angle of energy from the prism; (2) expression of E in terms of refractive index; and (3) the obtaining of the indices or their expression as a function of λ. The first step requires at least two constants, one for angular equivalent of the screw and another for the particular Littrow-mirror orientation for $D=0$. The second requires knowledge of the refracting angle of the prism and its orientation with respect to the incident energy. The adequate expression of index as a function of wavelength usually requires four constants.

Lithium fluoride is preferable to calcium fluoride for use in prismatic form in the infrared to about 6μ because of its much greater dispersion. Synthetic lithium fluoride has been available for 20 years or more, and data on its indices of refraction have been published by Littman [1][2] for the visible and to 3μ; also by Hohls [2] to 3 and 4 decimals from 0.546 to 12μ. The 4-decimal indices by Gyulai [3] are chiefly for the ultraviolet region.

Formerly, the making of lithium fluoride was carried out completely in an atmosphere of air, whereas more recently a vacuum technique has been used, with a concomitant change in some of the optical, mechanical, and thermal properties of the product [4]. Thus some question has arisen regarding possible differences in the refractivity of various specimens of this material. No final answer should be made at present, but it can be said that only small variations in the fifth decimal of refractive index have been found at this Bureau among six or more prisms of synthetic lithium fluoride that have been carefully measured in the visible region, including one that has been on hand for about 12 years and used as a secondary standard.

On this NBS standard prism refractive-index measurements were made in the visible region for five wavelengths in four groups of experiments at temperatures that averaged 28.6°, 31.6°, 52.6°, and 59.9° C. On two other prisms, at seven wavelengths, dispersion measurements were made at temperatures near 25° C and represented by a four-constant Ketteler-Helmholtz dispersion formula with the parameters adjusted by least squares. These prisms were lower in index than the above mentioned standard by only 1.4×10^{-5}, and no differences in dispersion could be detected. From these data, by the process of graduation, table 1 of the average refractivity of lithium fluoride was prepared at 5-deg intervals from 20° to 60° C at intervals of 100 A in wavelength from 4000 to 7000 A.

These data in the visible region yield tentative values for the temperature coefficient of refractive index. The average found for data on five wavelengths and for a very small temperature interval near 30° C is -16.4×10^{-6} per 1° C, and similarly for a small temperature interval near 56° the average is -17.5×10^{-6}. Some of the same data yield an average value of -16.3×10^{-6} for the interval 31.6° to 52.6° C. For red light the values are somewhat larger than for violet, perhaps 5 percent larger. These values agree well with the coefficient of -16.67×10^{-6} published by Littman, and are plotted in figure 1 for comparison with similar values reported by Radhakrishnan [5].

The coefficients in the right-hand columns of table 1 were obtained from differences smoothed in such manner that the results approximate average values for the derivatives. The change in index with wavelength applies almost equally well at all listed temperatures. The change with temperature is more nearly valid at the mid temperature of 40° C.

The refractive-index measurements herein reported for the infrared region were made on a prism with faces about 6 by 8 cm and a refracting angle, A, of 72°2′24″ as ground and polished by the Perkin-Elmer Corp. of Glenbrook, Conn. The first work was done on a minimum-deviation spectrometer for several wavelengths in the visible region in order to compare its refractivity with that of lithium fluoride prisms that had been previously measured. These index values proved to be intermediate with respect to, and closely an average of, those previously determined for other specimens. They are expressed with approximately sixth-decimal-place precision in

[1] The work described in this paper was carried out in part under the sponsorship of the U. S. Air Force at the National Bureau of Standards. A report thereon was presented on October 27, 1950 at the Cleveland meeting of the American Optical Society.

[2] Figures in brackets indicate the literature references at the end of this paper.

TABLE 1. *Refractivity of lithuim fluoride at various temperatures in the visible region*

[The column for 25° C was computed as $n^2=1.9261500-0.00705034\lambda^2+\frac{0.00471433}{\lambda^2-0.0094731}$]

Wave length	Temperature, °C									$-\frac{dn}{d\lambda}/\lambda$	$-\frac{dn}{dt}/°C$
	20°	25°	30°	35°	40°	45°	50°	55°	60°		
μ											
4000	1.398772	1.398696	1.398619	1.398541	1.398462	1.398382	1.398301	1.398219	1.398136	6.18×10^{-4}	15.8×10^{-6}
4100	1.398180	1.398104	1.398027	1.397949	1.397870	1.397790	1.397709	1.397627	1.397544	5.72	15.8
4200	1.397630	1.397554	1.397477	1.397399	1.397320	1.397240	1.397159	1.397077	1.396994	5.31	15.9
4300	1.397118	1.397042	1.396965	1.396887	1.396808	1.396728	1.396647	1.396565	1.396482	4.94	15.9
4400	1.396642	1.396566	1.396489	1.396411	1.396332	1.396252	1.396171	1.396089	1.396006	4.61	16.0
4500	1.396197	1.396121	1.396044	1.395966	1.395887	1.395807	1.395726	1.395644	1.395561	4.31	16.0
4600	1.395780	1.395704	1.395627	1.395549	1.395470	1.395390	1.395309	1.395227	1.395144	4.04	16.1
4700	1.395389	1.395313	1.395236	1.395158	1.395079	1.394999	1.394918	1.394836	1.394753	3.79	16.1
4800	1.395023	1.394946	1.394868	1.394789	1.394709	1.394628	1.394546	1.394463	1.394379	3.57	16.2
4900	1.394677	1.394600	1.394522	1.394443	1.394363	1.394282	1.394200	1.394117	1.394033	3.37	16.2
5000	1.394350	1.394273	1.394195	1.394116	1.394036	1.393955	1.393873	1.393790	1.393706	3.19	16.3
5100	1.394041	1.393964	1.393886	1.393807	1.393727	1.393646	1.393564	1.393481	1.393397	3.02	16.3
5200	1.393749	1.393672	1.393594	1.393515	1.393435	1.393354	1.393272	1.393189	1.393105	2.86	16.3
5300	1.393472	1.393395	1.393317	1.393238	1.393158	1.393077	1.392995	1.392912	1.392828	2.71	16.4
5400	1.393208	1.393131	1.393053	1.392974	1.392894	1.392813	1.392731	1.392648	1.392564	2.58	16.4
5500	1.392958	1.392880	1.392801	1.392721	1.392640	1.392558	1.392475	1.392391	1.392306	2.46	16.5
5600	1.392719	1.392641	1.392562	1.392482	1.392401	1.392319	1.392236	1.392152	1.392067	2.34	16.5
5700	1.392491	1.392413	1.392334	1.392254	1.392173	1.392091	1.392008	1.391924	1.391839	2.23	16.5
5800	1.392273	1.392195	1.392116	1.392036	1.391955	1.391873	1.391790	1.391706	1.391621	2.13	16.6
5900	1.392065	1.391987	1.391908	1.391828	1.391747	1.391665	1.391582	1.391498	1.391413	2.04	16.6
6000	1.391865	1.391787	1.391708	1.391628	1.391547	1.391465	1.391382	1.391298	1.391213	1.96	16.6
6100	1.391673	1.391595	1.391516	1.391436	1.391355	1.391273	1.391190	1.391106	1.391021	1.89	16.6
6200	1.391489	1.391410	1.391330	1.391249	1.391167	1.391084	1.391000	1.390915	1.390829	1.82	16.7
6300	1.391312	1.391233	1.391153	1.391072	1.390990	1.390907	1.390823	1.390738	1.390652	1.75	16.7
6400	1.391141	1.391062	1.390982	1.390901	1.390819	1.390736	1.390652	1.390567	1.390481	1.68	16.7
6500	1.390976	1.390897	1.390817	1.390736	1.390654	1.390571	1.390487	1.390402	1.390316	1.62	16.7
6600	1.390817	1.390738	1.390658	1.390577	1.390495	1.390412	1.390328	1.390243	1.390157	1.57	16.7
6700	1.390663	1.390584	1.390504	1.390423	1.390341	1.390258	1.390174	1.390089	1.390003	1.52	16.7
6800	1.390513	1.390434	1.390354	1.390273	1.390191	1.390108	1.390024	1.389939	1.389853	1.47	16.7
6900	1.390370	1.390289	1.390209	1.390127	1.390044	1.389960	1.389875	1.389789	1.389702	1.43	16.7
7000	1.390229	1.390149	1.390068	1.389986	1.389903	1.389819	1.389734	1.389648	1.389561	1.39	16.7

the visible for a temperature of 23.6° C by the formula

$$n^2=1.9262231-0.00705034\lambda^2+\frac{0.00471433}{\lambda^2-0.0094731},$$

where, as usual in such equations, λ is to be expressed in microns.

This formula was used in computing the index 1.39102 of this prism for the cadmium red line, $\lambda=6438$ A, which is the basic reference point in the index measurements for the infrared region. Also, to insure that the slope of the dispersion curve would be well controlled at the short-wavelength end, the indices computed by this formula were used as "observed" values for seven other wavelengths in the visible region.

The method used for index measurements is the same as that already described for measurements on crystals of thallium bromide-iodide [6] and on a crystal of silver chloride [7]. That is, briefly, the prism was installed with the aid of an auxiliary prism at an angle of known incidence on an infrared spectrometer equipped with a Littrow mirror; a thermocouple was used and a measure of the received energy was recorded on a Speedomax potentiometer. The incidence angle, i, for which the prism was installed was 54°51′28″, which corresponds fairly well to $(A+D)/2$ at minimum deviation for the cadmium red line. By ray tracing, one computes an angle of emer-

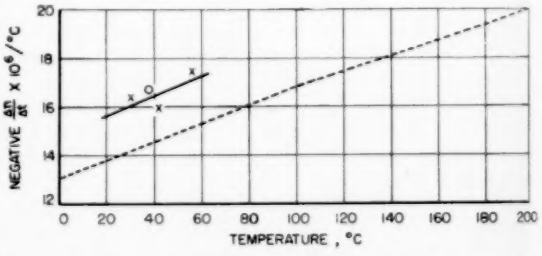

FIGURE 1. *Temperature coefficient of refractivity of lithium fluoride (averages for visible region).*

For red light the values are somewhat larger than for violet, perhaps 5 percent larger. X, NBS; O, Littman (n_D); ------, Radhakrishnan.

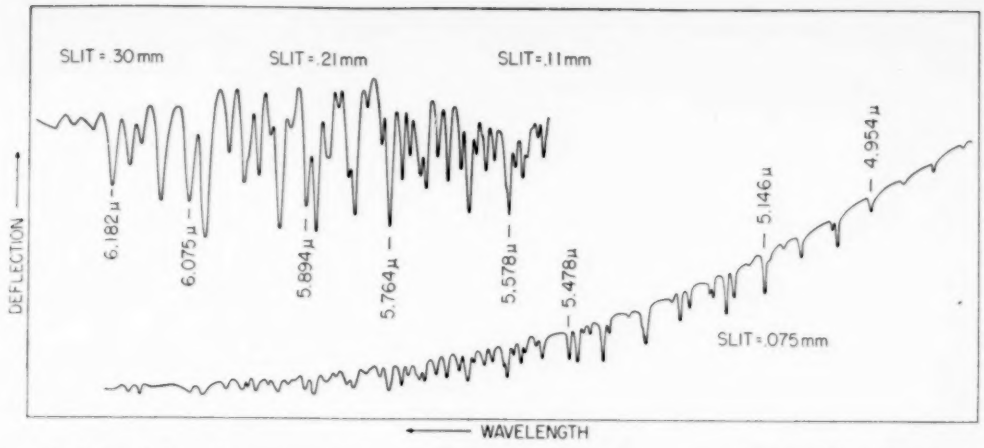

FIGURE 2. *Infrared absorption spectrum of water vapor in the 5- to 6-μ region as measured with a lithium fluoride prism.*
The upper curve was obtained with automatic slit control for constant radiant energy. The lower curve was recorded with a constant slit.

gence, E_0, of 54°54′55″ for the initial conditions at $\lambda = 6438$ A.

The angular increments for the Littrow mirror, between the initially and the subsequently recorded energy maxima and minima, when applied to the initial E_0 give approximate values of E_λ for use in computing indices by the equations

$$\left.\begin{aligned} \cot i' &= \cot A + \frac{\sin E}{\sin A \sin i'} \\ \text{and} \\ n &= \sin i(1+\cot^2 i')^{1/2}. \end{aligned}\right\} \quad (1)$$

If the entrance and exit slits of the infrared spectrometer subtend an angle at the collimating mirror, it is necessary to make appropriate correction for the fact that one-half of the corresponding angle at the face of the Littrow mirror is not constant as wavelength varies. The index-of-refraction measurements were made in the infrared region by the use of a Perkin-Elmer infrared spectrometer. The prism remains stationary, and different regions of the spectrum are obtained by rotation of a Littrow mirror. A new-type thermocouple, developed by the Perkin-Elmer Corp., was used as the detector. It was used with both continuous and modulated radiant energy. Final drum readings, however, were obtained by slow manual control of the drum near the maxima of deflections, in order to avoid errors caused by small time lags in the response of the pen to changes in energy at the detector.

The wavelengths used in the index measurements in the visible and near infrared region were obtained by the use of the radiant energy from an FH–4 mercury-cadmium lamp. Several lines from this source in the visible and near infrared region are sufficiently intense that they may serve as precise standards for determining the relationship of the Littrow mirror angle and wavelength. In table 2 are listed 15 lines of the spectrum of Hg and Cd that have been used for the measurement of the indices of refraction of LiF.

In the region from 2.7 to 6 μ nine absorption bands of methanol, polystyrene, carbon dioxide, and water vapor were used for the standard wavelengths. The wavelengths of the two absorption bands of methanol were taken from the work of Borden and Barker [8]. The values of indices of refraction experimentally determined by the use of 2.7144- and 4.866-μ absorptions of methanol differ by $+7\times10^{-5}$ and -9×10^{-5} from the adjusted values obtained by the use of the empirical equation (2). There may be a small error in the grating determinations of these two wavelengths.

The atmospheric water vapor absorption spectrum was recorded from 4.9 to 6.2 μ. The wavelengths of the lines were selected from the measurements of Plyler and Sleator [9] and H. H. Nielsen [10]. While there is some doubt as to the correctness of the wavelengths to the fourth decimal place, the accuracy is sufficient for the determination of these indices of refraction of LiF in the 5- to 6-μ region. In figure 2 there is shown a recording of the water vapor spectrum when medium amplification is used.

The upper curve was obtained when the automatic slit control was in operation, thus making it possible to hold the radiant energy almost constant. Because of the absorption of the prism material and its increase in dispersion with wavelength, it was necessary to increase the slit width by a factor of 3 in going from 5.50 to 6.2μ. The lower curve represents the recording of the water vapor spectrum with a constant slit. The wavelengths of some of the lines are marked on the graph. The numerical values are expressed in microns and are given to the third decimal place.

The refractive indices as computed from the observed data by eq (1) were adjusted by least squares through determination of the parameters of a four-

TABLE 2. *Observed and computed data on index of refraction of* LiF (23.6° C)

Wavelength	Source or absorbing medium	Computed index	(o−c)	Drum reading in turns×100	
				Observed	Computed
μ					
0.4047	Hg	1.39842	+2×10⁻⁴		
.4358	Hg	1.39678	0		
.4678	Cd	1.39542	0		
.4800	Cd	1.39497	0		
.5086	Cd	1.39403	0	2271.4	2271.1
.5461	Hg	1.39300	0	2244.5	2243.6
.5780	Hg yellow (mean)	1.39226	0	2223.8	2224.0
.6438	Cd	1.39101	+1	2191.8	2191.1
1.0140	Hg	1.38700	−1	2085.5	2085.4
1.1287	Hg	1.38609	+1	2062.3	2062.0
1.5295	Hg	1.38286	0	1980.9	1980.9
1.8131	Hg	1.38052	−1	1917.5	1917.9
1.9701	Hg	1.37904	0	1879.6	1880.0
2.2493	Hg	1.37616	−3	1805.0	1806.2
2.3254	Hg	1.37531	−1	1783.9	1784.6
2.7144	CH_3OH	1.37056	[a]+7	1665.5	1666.3
3.2432	Polystyrene	1.36291	−1	1472.0	1473.6
3.5078	do	1.35854	+2	1365.1	1365.9
4.258	CO_2 (center)	1.34410	+3	1015.1	1016.2
4.866	CH_3OH	1.33001	[a]−9	680.2	681.4
5.1456	H_2O vapor	1.32277	0	515.0	515.5
5.4638	do	1.31392	+2	313.5	312.5
5.7637	do	1.30496	−2	110.7	109.3
5.8944	do	1.30086	+3	20.5	17.1

[a] These larger residuals, +7 and −9×10⁻³, both for the wavelengths defined by absorption of methanol, would vanish if the wavelengths were in error by +0.0056 and −0.0036 μ, respectively.

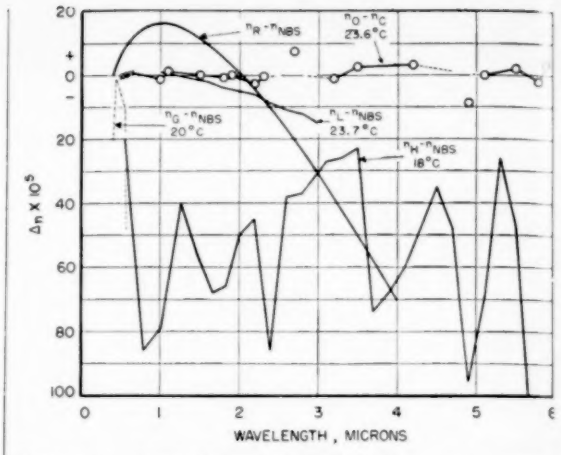

FIGURE 3. *Comparison of observed (circles) and computed* ($\Delta n=0$) *indices of refraction of lithium fluoride with previously reported values.*

The indices computed at NBS by formula (2) of this paper have been subtracted from all others. No allowances for effect of temperature have been made, but differences caused by temperature are less than 10×10⁻⁵. Recent values of index by Durie (J. Opt. Soc. Am. **40**, 879 (1950)) and values that he quotes by Kessler are in good agreement with NBS values. n_G, Gyulai; n_H, Hohls; n_L, Littman; n_O, observed at NBS; n_C, computed at NBS; n_R, Radhakrishnan.

constant Kettler-Helmholtz dispersion formula as limited to one infrared and one ultraviolet Sellmeier term. The result is

$$n^2=7.0537595-\frac{4091.74975}{797.8925296-\lambda^2}+\frac{0.00492029}{\lambda^2-0.0053160}, \quad (2)$$

which is preferable in computations to its equivalent form

$$n^2=1-\frac{5.1281966\lambda^2}{797.8925296-\lambda^2}+\frac{0.92556295\lambda^2}{\lambda^2-0.0053160}. \quad (3)$$

The agreement between the indices computed by this dispersion formula and those observed is excellent, as shown by the residuals tabulated in table 2 except for the two instances in which methanol was used. We think that the wavelengths for these two observations are slightly inaccurate and suggest that 2.7088 and 4.870 μ are better values than those listed in table 2 and used in computing the index data.

Figure 3 shows the results of this work on an infrared spectrometer compared with previously reported values of the index of lithium fluoride, the present observed values being represented by circles and the computed values by the line $\Delta n=0$. Table 3 expresses these computed values in a form for convenient interpolation at any value of λ from 0.5 to 6.0μ.

The precise fitting of these data by formula (2) is, we think, the first accurate representation of index of a solid by four constants over a comparable spectral range (the absolute term of formula (2) being 1 plus functions of the four constants). The five-constant form

$$n^2=b^2-\frac{M}{L^2-\lambda^2}+\frac{m}{\lambda^2-l^2} \quad (4)$$

where b^2 is independent of the other parameters, and also the five-constant form

$$n^2=a^2-K\lambda^2-p\lambda^4+\frac{m}{\lambda^2-l^2}, \quad (5)$$

(which latter we successfully used for KRS–5 and for Ag Cl) were each tried (in simultaneous solutions) for lithium fluoride, but neither seemed more suitable than the four-constant formula with only the two Sellmeier terms. However, this comparison of formulas is not definite because a least squares adjustment was made only on the one last named.

For KRS–5 it seems that the inferiority of the four-constant formula may be said to occur because one fictitious infrared band cannot represent the band near 117μ and also the other near 152μ as well as can be done by three terms in the expansion of each of their respective Sellmeier terms. On the other hand, for lithium fluoride, the band near 16μ is so comparatively unimportant that its effect and also that of the band near 23μ can be effectively represented as one. That the effective location (28 μ) of this one infrared band is not in this case intermediate between the actual locations mentioned is of no particular import. These dispersion equa-

TABLE 3. *Refractive indices of lithium fluoride at 23.6° C*

(Computed by formula 2)

Wavelength	Tenths of microns									
	0	1	2	3	4	5	6	7	8	9
μ										
0						1.39430	1.39181	1.39017	1.38896	1.38797
1	1.38711	1.38631	1.38554	1.38477	1.38400	1.38320	1.38238	1.38153	1.38064	1.37971
2	1.37875	1.37774	1.37669	1.37560	1.37446	1.37327	1.37203	1.37075	1.36942	1.36804
3	1.36660	1.36512	1.36359	1.36201	1.36037	1.35868	1.35693	1.35514	1.35329	1.35138
4	1.34942	1.34740	1.34533	1.34319	1.34100	1.33875	1.33645	1.33408	1.33165	1.32916
5	1.32661	1.32399	1.32131	1.31856	1.31575	1.31287	1.30993	1.30692	1.30384	1.30068
6	1.29745									

tions have an approximately theoretical basis, but caution is necessary in quantitative reasoning based on their use. Whenever parameters are determined by purely statistical methods their values must to some extent perform functions other than those attributable to their particular simple interpretations unless the formula is in every way theoretically adequate.

The advantages of using an equation rather than a graphical method in the calibration of an infrared spectrometer are pointed out by McKinney and Friedel [11], but such procedure cannot be relied on for accuracy [12] unless the equation used is theoretically applicable or is thoroughly tested throughout the range of use.

In special cases, particularly for short intervals of wavelength, sufficiently remote from the resonance regions, one may find comparatively simple relationships between drum readings, D, and the wavelength, λ, but in general with any spectrometer and prism the safe procedure is (1) to express D in terms of the emergent angle, E, and (2) to find the proper relation between E_λ and the corresponding refractive index, n_λ.

For the spectrometer used in this investigation, the drum readings are nearly proportional to changes in emergent angle, and one may write

$$D=(E-e)/s, \tag{6}$$

where e is the particular value of E corresponding to $D=0$, and s is an average angular equivalent of the screw.

The values of E for use in computing D for any prism of known refracting angle A are obtainable from eq 1, which may be written

$$\sin E=\sin A\,(n^2-\sin^2 i)^{1/2}-\cos A \sin i, \tag{7}$$

where i, the angle of incidence, will be known by the manner of installation ($\sin i=n \sin A/2$ if set for minimum deviation) and values of the index, n, must be obtained from a table or computed from an appropriate equation expressing n as a function of λ.

For this spectrometer and our installation of this prism, after writing equations

$$sD_\lambda+e=E_\lambda \tag{8}$$

for all our observations and adjusting by the method of averages, we find $s=0.444259°$ and $e=45.1802°$ if the unit of D is 0.01 turn. For our prism,

$$\sin E=0.9512720\,(n^2-0.6686737)^{1/2}-0.2521473, \tag{9}$$

and we have by these means obtained the computed values of D as listed in table 2.

In figure 4, where the observed and computed values of D are compared, it will be noticed that the observed values of D are systematically low by about 0.01 turn of the drum, between D readings of 440 and 1980 (turns $\times 100$) and of opposite sign elsewhere. These slight differences are a result of two causes. First, the angular equivalent of the screw is only an average value. A calibration of the screw was used in the precise reduction of the observed data but is ignored in this subsequent proposal for computation of drum readings. Second, the small angular subtense of the entrant and emergent slits is also ignored in this computation of drum readings from the adjusted observed data.

Consequently, if the slits are in near angular coincidence, and if the screw advance (reasonably free from periodic error) is nearly proportional to

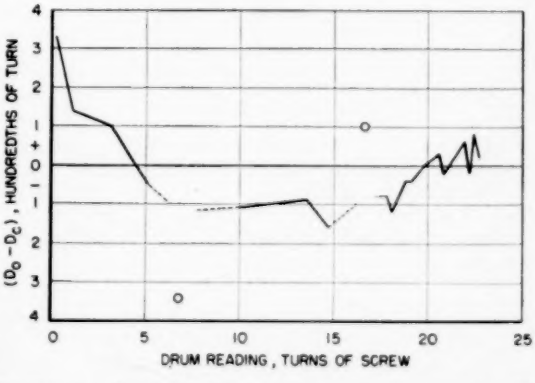

FIGURE 4. *Computed drum readings, D_c, compared with observed drum readings, D_o.*

The systematic difference occurs chiefly because the relation between drum reading and angular orientation of the mirror is not strictly linear. An error of one hundredth turn is approximately equivalent to an error of 4×10^{-5} in refractive index. (The two circles represent data based on the absorption bands of methanol for which there may be small errors in the reported values of wavelength.)

angle of orientation of the mirror, then a satisfactory calibration of the drum readings in terms of wavelength can be made, provided the drum readings are known for some 10 or 15 wavelengths for which the indices are reliably known or obtainable. In fact, if the drum readings can be observed with sufficient accuracy (higher than that needed in subsequent use of the spectrometer) then four or five wavelengths will in most cases be sufficient for the calibration (but with no remaining data to confirm the correctness of procedures).

For precision and reliability in expressing refractive index as a function of wavelength the Sellmeier dispersion equation

$$n^2=1+\frac{C\lambda^2}{\lambda^2-\lambda_c^2}+\frac{K\lambda^2}{\lambda^2-\lambda_r^2} \tag{10}$$

and its variants are unexcelled. It is best known in its slightly expanded form in which it is exactly equivalent to the Ketteler-Helmholtz equation having only four disposable parameters. For use in the most frequently encountered conditions, where the ultraviolet critical frequencies are proximate or more important than those in the infrared, one further expands the infrared term and writes

$$n^2=a^2-k\lambda^2-p\lambda^4+\frac{m}{\lambda^2-l^2}, \tag{11}$$

in which the term in λ^4 can often be omitted. Similarly, for the opposite case, where the infrared bands are near and predominant, one uses

$$n^2=a^2+\frac{b}{\lambda^2}+\frac{c}{\lambda^4}-\frac{M}{L^2-\lambda^2}, \tag{12}$$

but sometimes when both the ultraviolet and infrared effects are important it is advantageous, from the standpoint of high precision and accuracy, to keep the fractional remainders of both Sellmeier terms

$$n^2=a^2-\frac{M}{L^2-\lambda^2}+\frac{m}{\lambda^2-l^2}, \tag{13}$$

which may be only a four-constant formula (a^2 being a function of L^2, l^2, M, and m) or may be used with a^2 as an independent parameter.

The shorter the interval of wavelength, the more remote the critical frequencies, and the less precise the data, the greater are the chances for successfully using the fully expanded forms and those simpler forms written with the first power of n, such as

$$n=a+b\lambda^2+\frac{c}{\lambda^2}+\cdots \tag{14}$$

and the numerous empirical variants that are now and then proposed, tested, and found satisfactory under one or more of the three limitations just stated.

In selecting the best variant of these dispersion equations for use in any given instance, and in deciding how many constants should be used, it is usually advantageous to make simple simultaneous solutions before resorting to least squares. In such solutions it is advantageous to solve by systematic elimination and arrange the work so that the solution for $c-1$ constants can be made with very little additional work after a solution for c constants.[3] Likewise, in arranging the work for least squares solutions, one can so manage that most of the sums of the squares and cross products for the $c-1$ adjustment are obtained from among those computed for the c-constant adjustment, and then the normal equations can be so written and solved that little work will thereafter be necessary in obtaining the adjusted solution for $c-1$ constants.

Much of the effort made to substitute simplified dispersion equations for those of the Sellmeier type stems from the nonlinearity of the constants when Sellmeier terms are present. An objection is that direct least squares solutions are not possible, and the differential process of betterment of some approximate solution must be used. Much of this objection vanishes whenever it is advisable to make preliminary simultaneous solutions in the process of selection of a proper equation.

[1] H. Littman, Phys. Z. **41,** 468 (1940).
[2] H. W. Hohls, Ann. Physik [5] **29,** 433 (1937).
[3] Z. Gyulai, Z. Physik. **46,** 84 (1927).
[4] Combes and Ballard, J. Opt. Soc. Am. **38,** 1101A (1948); **39,** 635A (1949).
[5] T. Radharkrishnan, Proc. Indian Acad. Sci. **31A,** 226 (1950).
[6] Leroy W. Tilton, Earle K. Plyler, and Robert E. Stephens, J. Research NGS **43,** 81 (1949) RP2008.
[7] Leroy W. Tilton, Earle K. Plyler, and Robert E. Stephens, J. Opt. Soc. Am. **40,** 540 (1950).
[8] A. Borden and E. F. Barker, J. Chem. Phys. **6,** 553 (1938).
[9] E. K. Plyler and W. W. Sleator, Phys. Rev. **37,** 1493 (1931).
[10] H. H. Nielsen, Phys. Rev. **59,** 565 (1941).
[11] D. S. McKinney and R. A. Friedel, J. Opt. Soc. Am. **38,** 225 (1948).
[12] P. Porkington, J. Opt. Soc. Am. **40,** 481 (1950).

Washington, December 5, 1950.

[3] For example, the observational equation that is not to be used in the second solution should be written either first or last, but not among, the other $c-1$ observational equations, and one eliminates last, or solves first, for the constant not to be used later.

of Research of the National Bureau of Standards Vol. 47, No. 1, July 1951 Research Paper 2224

Preparation and Purification of Hydrogen Deuteride

Abraham Fookson, Philip Pomerantz, and Edwin H. Rich

In an investigation of the physical properties of hydrogen isotopes it was necessary to prepare 10 liters of high-purity hydrogen deuteride. This was done by reacting lithium aluminum hydride with deuterium oxide, giving hydrogen deuteride, in two batches, of 98.1- and 97.0-percent purity, respectively. This material was then fractionated at liquid hydrogen temperature in three batches, yielding the desired amount of hydrogen deuteride in a purity of 99.8 percent.

I. Introduction

In connection with a project on the measurement of precise physical properties of hydrogen isotopes, it was found necessary to prepare about 10 liters of high-purity hydrogen deuteride. This substance had been prepared for the first time by Scott and Brickwedde [2, 6] [1] by passing a mixture of hydrogen and deuterium over a heated metal filament, followed by fractionation at liquid hydrogen temperature. The equilibrium mixture after conversion contained about 50 mole percent hydrogen deuteride, as would be anticipated from the equation

$$H_2 + D_2 \rightleftarrows 2HD,$$

for which the equilibrium constant is about 4 [5].

Mixtures containing a high percentage of hydrogen deuteride have also been prepared by the action of deuterium oxide or water on various hydrides or deuterides, respectively. These methods give, in general, higher yields of hydrogen deuteride than is obtainable by equilibrating hydrogen and deuterium. Thus, Norton [5] treated boron hydride, B_2H_6, with deuterium gas to obtain the deuterated product, B_2D_6. The latter, when treated with 83-percent sulfuric acid, reacted with the water present to yield a mixture of hydrogen isotopes containing 85-percent hydrogen deuteride.

Similarly, lithium hydride with deuterium oxide has been used to prepare enriched hydrogen deuteride [1]. With sodium hydride, 87-percent hydrogen deuteride is obtainable [7], and with lithium aluminum hydride and deuterium oxide at 0° C, 99-percent hydrogen deuteride has been prepared [7].

For the purpose at hand, which requires hydrogen deuteride containing no more than several tenths percent of impurities, the last-mentioned reaction suggested itself as the most appropriate. In addition, to further purify the material obtained from this reaction, the fractionation technique of Scott and Brickwedde was employed.

II. Apparatus

The "crude" hydrogen deuteride was prepared in the apparatus described by Dibeler [4], shown in figure 1, except for the method employed for adding deuterium oxide. This reagent was added by means of a hypodermic syringe through a neoprene septum held in a fitting attached to the tapered joint, *C*. It was found that the septum could be punctured several times without danger of admitting air to the system. This rendered it possible to add the deuterium oxide in several portions, thereby diminishing the vigor of the reaction.

In addition to the method for adding deuterium oxide, the apparatus of Dibeler was further modified by the inclusion of a magnetic stirrer in the flask, *D*; this was operated by a rotating magnetic field outside the flask.

The crude hydrogen deuteride was distilled in the apparatus shown in figure 2. This is, with minor modifications, the apparatus used by Scott and Brickwedde [2, 6] in their work; this technique has not heretofore been published.[2] The still, *A*, consisted of a jacketed boiler of about 5-ml capacity and monel helix rectifying section, *B*, a re-entrant portion, *C*, which served as a cold-finger dephlegmating condenser, and a take-off tube, *E*, which was connected to the top of the still by means of a U-tube, *D*. The liquid was boiled by means of an electric heater, *F*, of constantan wire, the leads of which ran up the still and emerged through a vacuum-tight wax seal, *G*. The still was immersed in a Dewar flask of liquid hydrogen, the level of which was maintained at about the position indicated in the drawing. This Dewar was in turn immersed in another Dewar containing liquid nitrogen (not shown in the drawing).

The still was connected through a ground glass joint to a regulator valve, *H*, thence to a graduated Toepler pump, *J*, of 500-ml capacity. The regulator valve consisted of a length of thin drill rod brazed to an iron core. The drill rod fit with little clearance into a length of heavy-walled capillary tubing, and was movable within this capillary by means of the solenoid, *K*. By thereby varying the length of annular space between the drill rod and tube wall, the rate of flow of gas from the still to the Toepler pump was easily controlled. The pressure in the still was indicated by a manometer; by adjusting the levels in the Toepler pump by means of the mercury reservoir, *M*, to a pressure less than that shown on the manometer, gas could be caused to flow from

[1] Figures in brackets indicate the literature references at the end of this paper.

[2] Clusius and Starke [3] have recently described a similar procedure. This was developed subsequent to, and independent of, Scott and Brickwedde.

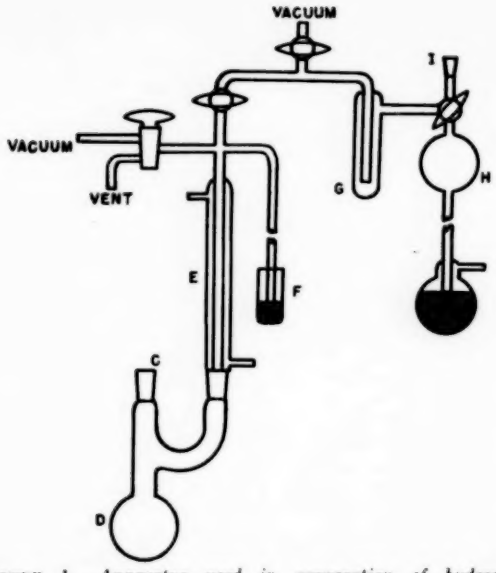

FIGURE 1. *Apparatus used in preparation of hydrogen deuteride.*
This drawing is identical with that reproduced in V. H. Dibeler's paper [4].

the still to the Toepler pump at a rate determined by the pressure difference and the setting of the regulator valve.

The Toepler pump was connected to a manifold, *N*, and from there to a mercury diffusion pump. The manifold was fitted with a number of ground glass joints, to which sample tubes, receiver bulbs, or bulbs containing charging stock could be attached. The general practice was to use the charging bulbs as receivers, after the material they contained had been added to the still and the residual gas removed by the vacuum pump.

III. Preparation of Hydrogen Deuteride

Since 10 liters of high-purity hydrogen deuteride were required, this could readily be obtained by distillation from 15 liters of crude prepared by the method of Wender, Friedel, and Orchin [7] (lithium aluminum hydride and deuterium oxide at 0° C). Generous allowance could then be made for losses incurred during handling and distillation.

Accordingly, 15 liters of hydrogen deutride were prepared in two runs, of 5 and 10 liters, respectively, of 98.1- and 97.0 percent purity. Since both batches were prepared in the same manner, only one will be described.

Into the Claisen flask, *D* (letters in this section refer to fig. 1), there was distilled about 150 ml of *n*-butyl ether, which had been previously dried and distilled from sodium. This was followed by 5.75 g of lithium aluminum hydride, the addition of this reagent being carried out in a nitrogen atmosphere. This quantity of lithium aluminum hydride was 40 percent in excess of that required to prepare 10 liters of gas, as calculated from the equation

$$LiAlH_4 + 4D_2O \rightarrow LiOD + Al(OD)_3 + 4HD.$$

The rubber septum-retaining device was fitted to the flask, at *C*, and the latter attached to the condenser, *E*.

The reaction mixture was frozen by means of a liquid nitrogen bath, the system evacuated, and the reaction mixture heated until it boiled under its own vapor pressure. After 1½ hours, the flask was again cooled with liquid nitrogen, and the system re-evacuated. By means of a hypodermic syringe, 5 ml of deuterium oxide (99.5-percent purity) was added to the solid reaction mixture. The latter was permitted to melt, and the gas evolved as a consequence of the ensuing reaction was collected in two 5-liter collecting bulbs attached to a manifold at *I*. During this operation, the mixture was being stirred continuously. Due to its low temperature, the outside of the flask quickly accumulated a layer of frost. Whenever this began to melt, the liquid nitrogen bath was applied; in this manner, the reaction temperature was kept at 0° C or lower.

During the course of the reaction, deuterium oxide was added two more times, to make a total of 18 ml of this reagent added. This quantity was about 150 percent excess over that required to make 10 liters of gas as calculated from the above equation. It was found desirable to use this excesss in order to prevent the reaction from becoming too sluggish as the reagents were consumed.

After each run was completed, a sample was withdrawn from the collecting bulb by means of the Toepler pump, and added to an evacuated sample tube. These samples were subsequently analyzed mass spectrographically to determine purity of the hydrogen deuteride. This was found to be 98.1 and 97.0 percent, respectively, for the two batches prepared.

IV. Distillation of Hydrogen Deuteride

Since about 935 ml of hydrogen deuteride gas at 20° C is equivalent to 1 ml of liquid at the boiling point of hydrogen, 20.34° K, one 5-liter bulbfull of the crude as prepared was used as a convenient charge for each distillation. The entire 15 liters was therefore distilled in three batches, of which only one will be described in detail.

The collecting bulb, *S* (letters in this section refer to fig. 2) containing the charge at atmospheric pressure was attached to the manifold, as shown, and the entire system evacuated by means of a mercury diffusion pump. The Dewar flasks surrounding the still were filled with liquid hydrogen and liquid nitrogen, stopcock *R* was closed, and stopcocks *U*, *O*, *P*, and *Q* opened; the regulator valve, *H*, was also opened wide by means of solenoid *K*. The contents of *S* were thereby permitted to flow into the still, where condensation occurred on the condenser surface, *C*. After about 1 hour, the jacketed portions of the still had cooled sufficiently to permit liquid to begin accumulating in the boiler. The Toepler pump was then used to withdraw the remain-

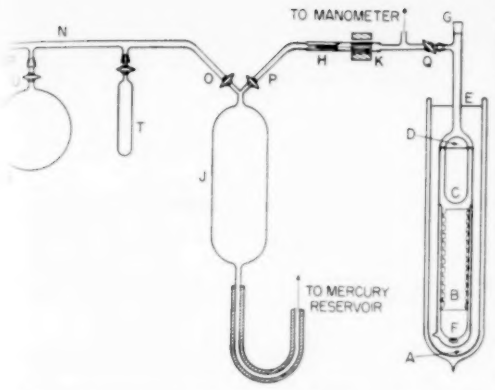

FIGURE 2. *Apparatus used for purification of hydrogen deuteride by fractional distillation.*

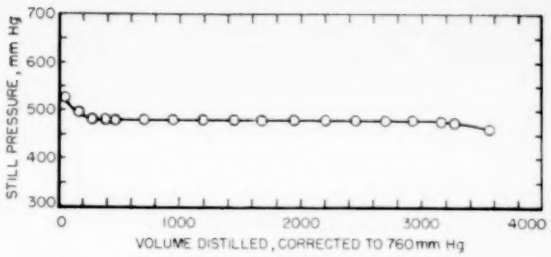

FIGURE 3. *Distillation curve of hydrogen deuteride.*

ing gas from S and to force it into the still; this last operation was continued until the pressure in the bulb as read on the manometer (stopcock Q closed), was 50 mm. It was not considered practical, with regard to the amount pumped for the time consumed, to attempt to reduce the pressure any lower.

Stopcock P was then closed, and a current of 71 ma permitted to flow through the heater, F. This current was controlled by means of a variable transformer; since the heater had a resistance of 27 ohms, it was equivalent to a heat input of 0.136 watt, or 0.033 cal/sec. The charge was thus caused to boil under total reflux, and, as the hydrogen present in the still was concentrated in the head as a consequence of fractionation, the pressure indicated on the manometer rose from an initial value of 447 to 531 mm. This rise occurred in about 45 min and did not increase thereafter. (Had there been sufficient hydrogen in the charge, the pressure at equilibrium should have been about atmospheric, or somewhat greater due to the fact that heat was being introduced.)

During this time, stopcocks R, U, and O were open, evacuating the Toepler pump and the bulb, S, for use as a receiver. Other, smaller, bulbs were also on the manifold, to be used as forerun and residue receivers; these are not shown in figure 2.

When equilibrium had been reached, as indicated by constancy of pressure, the regulator valve was adjusted to about half-open, stopcock O was closed, and stopcock P opened. The mercury levels in J and M were adjusted so that the pressure in the Toepler pump was about 160 mm less than that in the still. Take-off of distillate was thereby begun.

Each time the bulb J filled, stopcock P was closed, stopcock O was opened, and the contents of the Toepler pump were discharged into the appropriate receiver. The points at which fractions were to be cut were determined by the pressure indicated on the manometer. This pressure was approximately the equilibrium pressure of the distillate at the condenser temperature. Its absolute value was of little concern; what was of interest during the distillation was its variation. As can be seen from figure 3, the pressure-volume distilled curve for this distillation, the bulk of the distillate was removed at 478 to 480 mm. The pressure at any time was influenced by the level, E, of liquid hydrogen, the take-off rate at that moment, the ambient temperature, and the rate of boiling.

During the course of the distillation it was necessary to adjust the regulator valve several times and to determine the proper pressure in the Toepler pump, in order to achieve the desired reflux ratio. The heat of vaporization of hydrogen deuteride is 257 cal mole^{-1} at its boiling point, 22.54° K [8]. Since the heat input was 0.033 cal sec^{-1}, the boil-up rate was 185 ml min^{-1}, at 20° C and 760 mm. The take-off rate was observed by noting the amount withdrawn in the Toepler pump, and the time taken to withdraw it. From these figures, the reflux ratio was calculated. Because of the nature of the adjustments necessary to the operation of the still (as the mercury level in J fell, the leveling bulb, M, had to be periodically lowered to maintain approximately the same head), the reflux ratio was variable. During the withdrawal of the pure hydrogen deuteride fraction, this quantity ranged from 9.3:1 to 26.5:1, with an average of 13.7:1.

In the distillation described above, the charge was equivalent to 4,600 ml of gas, at 20° and 1 atm. There was distilled 3,580 ml, of which 390 ml was forerun, 2,850 ml pure hydrogen deuteride, and the rest holdup. The volumes of the fractions were obtained by summing the volumes of gas discharged from the Toepler pump and correcting to 20° C and 1 atm; these values are plotted on the distillation curve, figure 3. Upon completion of the distillation, the pressure in the bulb containing the pure hydrogen deuteride was measured on the manometer. The corrected volume as calculated from this figure was 3,080 ml. It is believed that this value is more reliable than that obtained by summing the Toepler pump fractions, since the latter undoubtedly contain cumulative errors.

The distillation was continued until the boiler and the lower coils of the fractionating section were dry. This left a holdup of about 1 liter of gas, since with such a small amount of material fractionation was impossible in this still. The forerun and holdup of the distillation were discarded.

A sample of the main fraction was analyzed by means of the mass spectrometer, and was shown to have a purity of 99.8-percent hydrogen deuteride, disregarding a small percentage of nitrogen. The remainder of the crude hydrogen deuteride was distilled in the manner described herein, and yielded material of about the same purity.

The authors express their appreciation to V. H. Dibeler, whose apparatus was used in the preparation of the crude hydrogen deuteride, and who performed the mass spectrographic analyses reported in this work. Thanks are also due R. B. Scott and W. E. Gifford, who assisted in the distillation described in this paper.

V. References

[1] H. Beutler, G. Brauer, and H. C. Jünger, Naturwissenschaften, **24,** 347 (1936).
[2] F. G. Brickwedde and R. B. Scott, Phys. Rev. **55,** 672 (1939).
[3] K. Cl, and K. Starke, Z. Naturforsch **4a**, 549 (1949).
[4] V. H. Dibeler, J. Research NBS **44,** 363 (1950) RP 2083.
[5] F. J. Norton, Science **111**, 202 (1950).
[6] R. B. Scott and F. G. Brickwedde, Phys. Rev. **48,** 483 (1935).
[7] I. Wender, R. A. Friedel, and M. Orchin, J. Am. Chem. Soc. **71,** 1140 (1949).
[8] H. W. Woolley, R. B. Scott, and F. G. Brickwedde, J. Research NBS **41**, 379 (1948) RP 1732.

WASHINGTON, November 29, 1950.

of Research of the National Bureau of Standards Vol. 47, No. 1, July 1951 Research Paper 2225

Linear Thermal Expansion of Artificial Graphites to 1,370° C [1]

Milton D. Burdick, Benson Zweig, and R. Eugene Moreland

In connection with creep studies of graphite, the need arose for accurate thermal expansion data on the three grades of graphite being investigated. Therefore, the linear thermal expansion of specimens of the three grades, cut in three directions with respect to the direction of extrusion or of molding pressure, was determined at various temperatures between 20° and 1,370° C in vacuum. Results show that specimens cut perpendicular to the extrusion direction or parallel to the direction of molding pressure had greater expansions than specimens cut parallel to the extrusion direction or perpendicular to the molding pressure. The data obtained are tabulated.

I. Introduction

In connection with creep studies of graphite, the need arose for accurate thermal expansion data within the range from 980° to 1,370° C on the particular graphites being studied. The data were needed primarily to correct for the thermal expansion that would result from changing the test temperature during elongation or creep tests of graphite in tension.

Because of the marked anisotropy of the graphite crystal,[2 3 4 5] any departure from a random orientation, as may result from the forming process, would be expected to have an effect on the thermal expansion in various directions with respect to the direction of forming. Specimens for creep testing were cut both perpendicular and parallel to the direction of extrusion or of molding pressure. Therefore, specimens for thermal expansion determinations were cut similarly to duplicate the crystal orientation of the creep specimens.

In order that the thermal expansion data might be of more general use, expansion values were obtained while heating from room temperature to 1,370° C, and contraction values were obtained during cooling from the maximum temperature to 400° C.

II. Materials

Three grades of artificial graphite were obtained in the form of blocks. For each grade, at least two specimens were made representing each of three orientations with respect to the direction of molding pressure or of extrusion. Additional specimens were used to check the reproducibility of the testing equipment and the uniformity of the block samples. The specimens were right cylinders machined to approximately ⅜ in. in diameter by 2 in. in length. Figure 1 shows how the specimens were cut from the blocks. The cross-sectional area of the blocks perpendicular to the direction of extrusion or of molding pressure represents the die opening for the extruded Grades A and B and the mold size for Grade C.

III. Apparatus and Methods of Test

The dilatometer used in determining the thermal expansion of graphite was similar to that used in an investigation of zirconia.[6] The dilatometer consisted of an alumina base into which three artificial sapphire rods were cemented to support the dial gage and a fourth sapphire rod centrally located to provide the mechanical linkage between the dial gage and the cylindrical graphite specimen, which rested on the alumina base. Dial gages reading either to 0.0001 in. or to 0.01 mm were used. Gage readings were corrected for the expansion of the sapphire. The values for the expansion of the sapphire were taken from a determination from room temperature to 1,000° C by the interferometer method and extrapolated to 1,400° C (see footnote 6). The specimen and the lower part of the dilatometer were heated in a wire-wound furnace, the inner winding of which consisted of platinum—20 percent rhodium and the outer winding of Chromel A. The furnace and the dilatometer fitted under the bell jar of a vacuum coating unit.[7] The shield clamped to the tripod below the dial gage never reached a temperature above 70° C during a test. The equipment is shown in figure 2.

The measurement of the temperature offered greater difficulty experimentally than did determinations of the change in length. The temperature near the specimen in the first 22 tests was measured with a

[1] This project was sponsored by the NEPA division of the Fairchild Engine and Airplane Corp.

[2] J. B. Nelson and D. P. Riley, The thermal expansion of graphites from 15°C to 800° C: Part I experimental; D. P. Riley, The thermal expansion of graphite: Part II, theoretical, Proc. Phys. Soc. London **57**, 477 (1945).

[3] Jacques Pierry, Physique etude sur la dilattaion du graphite Comp. Rend. L'Academie Des Sciences, Paris **223**, 501 (1946).

[4] Peter Hidnert, Thermal expansion of artificial graphite and carbon, J. Research NBS **13**, 37 (1934) RP693.

[5] Peter Hidnert and W. T. Sweeney, Thermal expansion of graphite, BS Tech. Pap. 335 (1927).

[6] R. F. Geller and P. J. Yavorsky, Effects of some oxide additions on the thermal length changes of zirconia, J. Research NBS **35**, 87 (1945) RP1662.

[7] Manufactured by Distillation Products Inc., Rochester, N. Y. Type LC1-500-3.

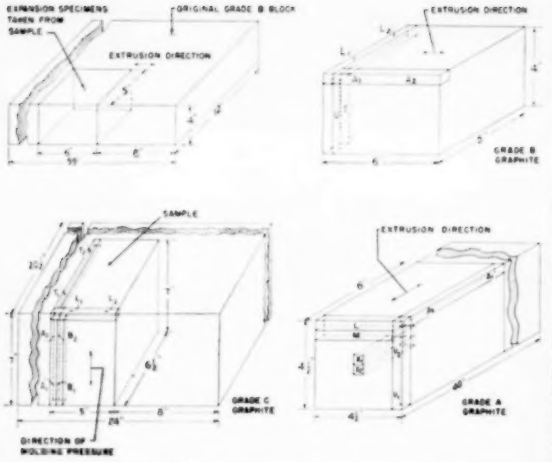

FIGURE 1. *Orientation of thermal expansion specimens in sample blocks and sizes of original blocks of graphite.*

30 gage Pt-Pt 10-percent Rh thermocouple in a closed-end protection tube extending radially through the furnace to the specimen. The protection tube was needed to prevent reaction between the graphite and the thermocouple metal. This reaction, the nature of which was not determined, had been observed in tests at temperatures above 980° C. In later tests, it was found that chromel-alumel thermocouples could be used without a protection tube. Although a new base-metal couple was used in each test, those removed were bright and could probably have been reused since subsequent calibration disclosed no deleterious effect.

Before making a test, the bell jar was evacuated to less than 0.05 mm of mercury by mechanical pumps. At the beginning of the test the oil diffusion pump was started. The vacuum attained by the combination of mechanical and oil diffusion pumps was usually less than 2×10^{-4} mm of mercury during the test. The heating rate was then adjusted to 3° C per minute, and measurements of expansion were obtained at 30° C intervals by observing dial readings through the glass wall of the bell jar. The maximum temperature of 1,370° C was held for 45 min after which the specimen was cooled to 400° C at the same rate of 3° C per minute. Air was admitted when the specimen had cooled to 400° C to permit reaching room temperature by the next morning.

For tests 1 through 22, for which a closed-end protection tube was used, the thermocouple had to be calibrated at the prescribed rate of heating to obtain true temperature readings. It was found that, because the protection tube could not be placed in the same position for each test, an error resulted in the temperature readings especially at the beginning of the tests. The effect of this error was reduced by averaging of the values.

The average deviation between expansion values determined under similar conditions for the same

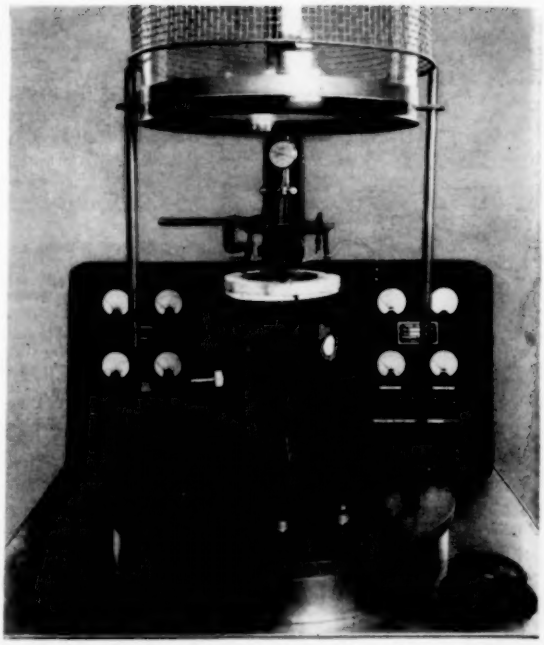

FIGURE 2. *Equipment used for determining thermal expansion of graphite.*

grade and orientation of graphite specimens was found to be ±0.02 percent (see table 3).

A test of the apparatus was made by measuring the expansion of a single crystal of periclase (MgO) to 1,370° C in vacuum and comparing the results with a previous dilatometer measurement on the same specimen heated to 1,700° C in air (see footnote 6). The comparison is given in table 1.

TABLE 1. *Comparison of the linear thermal expansion data obtained on the same MgO specimen in vacuum and in air*

Room temperature to—	Linear thermal expansion in—	
	Vacuum [a]	Air [b]
° C	*Percent*	*Percent*
100	0.09	0.06
200	.20	.16
300	.32	.30
400	.47	.43
500	.61	.58
600	.75	.71
700	.89	.87
800	1.04	1.01
900	1.19	1.17
1,000	1.32	1.32
1,100	1.47	1.48
1,200	1.61	1.64
1,300	1.77	1.81
1,370	1.88	1.93

[a] No loss in weight was found as a result of heating in a vacuum.
[b] Geller and Yavorsky (see footnote 6). No data on cooling were reported by these observers.

Another check of the apparatus was obtained by comparing the dilatometer measurements with those obtained by the interferometer method between

room temperature and 400° C. The agreement obtained with specimens of grade A graphite is given in table 2.

IV. Results and Discussion

The linear thermal expansions and contractions expressed in percentage for 100-deg intervals from room temperature to 1,370° C, and the average coefficients of expansion and contraction per degree centigrade over 200-deg intervals, are given in table 3 for each of the three grades of graphite and for the three directions with respect to forming pressure (one parallel and two perpendicular). Values are given for at least two specimens representing each grade and orientation with respect to the direction of forming. Also, tests were repeated twice for specimen X_1 of grade A and once for specimen A_1 of grade B to show the reproducibility. Moreover, specimens X_1 and X_2 of grade A were taken from the inside of the block (see fig. 1) to compare with specimens A_1 and A_2 of grade A from the outside face of the block, both pairs of specimens cut parallel to the extrusion direction. The average of the linear thermal expansion values from table 3 are shown in figure 3. Average coefficients of expansion computed from the curves in figure 3 are given for intervals of 200-deg centigrade in table 4.

TABLE 2. *Comparison of expansions of grade A graphite determined by the interferometer and the dilatometer methods*

Orientation	Temperature range	Coefficients of linear thermal expansion		Linear thermal expansion	
		Interferometer	Dilatometer [a]	Interferometer	Dilatometer
	°C			*Percent*	*Percent*
Parallel to extrusion direction A [b]	RT [c] to 100	0.7×10⁻⁶	0.8×10⁻⁶		
	100 to 200	1.1	1.6		
	200 to 300	1.5	2.1		
	300 to 400	2.0	2.6		
	RT to 400			0.05	[b] 0.07 (A_1) .06 (A_2)
Perpendicular to extrusion direction L, M [b]	RT to 100	2.8	3.6		
	100 to 200	3.8	4.7		
	200 to 300	4.3	5.0		
	300 to 400	4.4	5.0		
	RT to 400			0.15	0.18 (L) .17 (M)
Perpendicular to extrusion direction U [b]	RT to 100	4.0	4.5		
	100 to 200	5.2	4.8		
	200 to 300	5.5	4.8		
	300 to 400	5.6	5.5		
	RT to 400			0.19	0.19 (U_1) .18 (U_2)

[a] Average of 2 companion specimens.
[b] See orientation diagram, fig. 1.
[c] RT = room temperature.

The differences in the expansions perpendicular and parallel to the direction of forming, for the three grades of graphite, are attributed to the different degree of preferred orientation, and to the marked anisotropy of the graphite crystal. The positions of the crystals depend on the method of forming. With the extruded graphites, the C axes of the hexagonal crystal plates tend to be alined perpendicular to the sides of the die, this tendency being a function of the size of the grains, the density of the mass, the extrusion pressure, the size and shape of the die, etc. In the case of the molded graphite, the crystal plates tend to be alined with the C axes parallel to the sides of the mold. Since the direction of greatest expansion of the graphite crystal is parallel to the C axis (i. e., perpendicular to the basal cleavage), the specimens cut parallel to the direction of molding pressure and perpendicular to the extrusion direction showed the greater expansion. As is generally the case with clay bodies, the crystals are preferentially oriented to a greater degree by extrusion than by molding. This accounts for the fact that in the case of graphites A and B the difference in expansion between specimens cut parallel and perpendicular to the extrusion direction is greater than the difference for comparable specimens of molded graphite C.

In the case of grade B graphite, the two pairs of specimens (U, T and L_1, L_2), which were cut perpendicular to each other and to the direction of forming, differed significantly from each other in their linear thermal expansion. For comparable specimens of grades A and C, this difference was not significant. The large and rectangular die opening used in extruding the grade B graphite may have allowed a different degree of orientation of crystals between die faces than the relatively small square opening used for extruding grade A. This same explanation may account for the uniformity of expansion values for specimens of grade A cut parallel to the extrusion direction, despite the fact that specimens X_1 and X_2 were cut from the center of the block. Note in table 3 the uniformity of the expansion values for various specimens of grade A graphite cut perpendicular and parallel to the extrusion direction for tests in which the temperature was measured in the same manner, that is, with the protected thermocouple or with the bare thermocouple.

The results of the determinations for indicating the reproducibility are given in table 5. Comparison shows that the values do not deviate materially. Note that in test 44 the temperature was maintained for at least 1 hour at each of 10 temperatures. This

TABLE 3. *Linear thermal expansion and contraction data of oriented specimens from three grades of graphite*

Orientation [a]	Specimen	Test No. [b]	Linear thermal expansion and contraction from room temperature to—													
			100° C	200° C	300° C	400° C	500° C	600° C	700° C	800° C	900° C	1,000° C	1,100° C	1,200° C	1,300° C	1,370° C [c]
Grade A (extruded)																
			Percent	*Percent*	*Percent*	*Percent*	*Percent*	*Percent*	*Percent*	*Percent*	*Percent*	*Percent*	*Percent*	*Percent*	*Percent*	*Percent*
Parallel	A_1	20	0.01	0.02	0.04	[d] 0.07(.05)	0.10(.07)	0.13(.10)	0.16(.13)	0.19(.16)	0.22(.19)	0.25(.22)	0.28(.26)	0.32(.30)	0.36(.34)	0.38(.38)
	A_2	21	.00	.02	.04	.06(.04)	.09(.07)	.12(.10)	.15(.12)	.18(.16)	.21(.19)	.25(.22)	.28(.26)	.32(.30)	.37(.34)	.40(.39)
	X_1	18	.01	.04	.06	.09(.07)	.11(.10)	.14(.12)	.17(.15)	.20(.18)	.23(.21)	.27(.24)	.30(.28)	.34(.32)	.37(.36)	.40(.40)
		43	.00	.01	.03	.05(.00)	.07(.02)	.10(.04)	.13(.06)	.16(.09)	.19(.12)	.22(.15)	.2[illegible](.19)	.28(.24)	.32(.29)	.35(.32)
		[e] 44	.00	.01	.03	.05	.07	.09	.12	.15	.18	.21	.24	.27	.31	.33
	X_2	19	.00	.02	.04	.07(.05)	.10(.07)	.13(.10)	.16(.13)	.18(.16)	.22(.19)	.25(.22)	.28(.26)	.31(.29)	.35(.33)	.38(.37)
Perpendicular	L	14	.03	.08	.13	.18(.14)	.24(.19)	.29(.24)	.34(.29)	.40(.35)	.45(.41)	.51(.47)	.57(.53)	.63(.61)	.70(.69)	.74(.74)
	M	17	.02	.07	.12	.17(.13)	.22(.19)	.28(.24)	.34(.30)	.39(.35)	.45(.41)	.51(.48)	.57(.54)	.64(.61)	.70(.69)	.75(.74)
Do	U_1	15	.03	.08	.13	.19(.17)	.24(.22)	.30(.28)	.36(.33)	.42(.39)	.48(.45)	.54(.51)	.61(.57)	.67(.65)	.74(.72)	.79(.78)
	U_2	16	.04	.08	.13	.18(.15)	.24(.20)	.29(.25)	.34(.31)	.40(.36)	.46(.42)	.51(.48)	.58(.54)	.64(.61)	.70(.68)	.74(.74)
Grade B (extruded)																
Parallel	A_1	28	0.01	0.03	0.05	0.07(.02)	0.09(.04)	0.11(.06)	0.14(.09)	0.16(.11)	0.19(.14)	0.22(.17)	0.25(.20)	0.28(.24)	0.32(.29)	0.34(.33)
		40	.01	.03	.05	.07	.10(.00)	.12(.02)	.15(.05)	.17(.07)	.20(.10)	.23(.15)	.26(.16)	.29(.21)	.32(.26)	.34(.30)
	A_2	33	.01	.03	.06	.08(.02)	.11(.04)	.14(.07)	.16(.10)	.20(.12)	.22(.16)	.26(.20)	.29(.24)	.32(.28)	.37(.34)	.40(.38)
Perpendicular	L_1	34	.02	.06	.09	.13(.06)	.17(.10)	.21(.14)	.25(.18)	.29(.22)	.34(.26)	.38(.31)	.43(.37)	.48(.43)	.53(.50)	.57(.55)
	L_2	35	.02	.06	.10	.14(.07)	.18(.10)	.22(.14)	.26(.18)	.31(.22)	.36(.28)	.41(.33)	.46(.39)	.51(.46)	.56(.54)	.61(.60)
Do	U	36	.03	.08	.13	.17(.07)	.21(.31)	.26(.16)	.31(.20)	.36(.26)	.41(.32)	.46(.38)	.52(.45)	.58(.52)	.64(.60)	.68(.66)
	T	37	.03	.07	.12	.16(.09)	.21(.13)	.25(.17)	.30(.21)	.35(.26)	.40(.32)	.45(.38)	.51(.44)	.57(.52)	.63(.60)	.68(.66)
Grade C (molded)																
Parallel	A_1	22	0.03	0.08	0.14	0.19(.16)	0.24(.22)	0.30(.27)	0.35(.32)	0.41(.38)	0.47(.44)	0.53(.50)	0.59(.56)	0.65(.62)	0.72(.70)	0.76(.76)
	A_2	23	.03	.07	.11	.16(.07)	.20(.12)	.24(.16)	.29(.21)	.34(.26)	.39(.31)	.44(.36)	.49(.42)	.54(.48)	.59(.55)	[f].60(.60)
	B_1	29	.04	.08	.12	.17(.12)	.22(.16)	.26(.20)	.31(.25)	.36(.31)	.41(.36)	.47(.40)	.52(.47)	.58(.53)	.63(.61)	.68(.66)
	B_2	32	.03	.08	.12	.17(.12)	.22(.17)	.27(.21)	.32(.26)	.38(.32)	.43(.37)	.48(.44)	.54(.50)	60(.56)	.65(.63)	.70(.68)
Perpendicular	T_1	24	.02	.05	.09	.12(.08)	.16(.11)	.2[illegible](.14)	.24(.18)	.28(.22)	.33(.25)	.37(.30)	.42(.34)	.46(.40)	.50(.46)	[f].54(.54)
	T_2	25	.03	.06	.09	.12(.06)	.16(.10)	.19(.13)	.23(.17)	.27(.21)	.31(.25)	.35(.29)	.39(.33)	.44(.39)	.49(.45)	[f].52(.50)
Do	L_1	26	.02	.05	.08	.11(.06)	.15(.09)	.18(.12)	.22(.16)	.26(.20)	.30(.24)	.34(.29)	.39(.33)	.44(.38)	.48(.44)	[f].52(.50)
	L_2	27	.02	.05	.09	.12(.08)	.15(.13)	.19(.16)	.23(.19)	.26(.22)	.30(.25)	.35(.29)	.39(.33)	.44(.39)	.48(.44)	[f].52(.50)

TABLE 3. *Linear thermal expansion and contraction data of oriented specimens from three grades of graphite*—Continued

Orientation [a]	Specimen	Test No. [b]	Coefficients of linear thermal expansion— Room temperature to 200° C	200° to 400° C	400° to 600° C	600° to 800° C	800° to 1,000° C	1,000° to 1,200° C	1,200° to 1,370° C [c]
Grade A (extruded)									
Parallel	A_1	20	1.3×10^{-6}	2.4×10^{-6}	$2.8(2.8)\times10^{-6}$	$2.9(3.0)\times10^{-6}$	$3.2(3.7)\times10^{-6}$	$3.5(3.6)\times10^{-6}$	$3.7(4.5)\times10^{-6}$
	A_2	21	1.1	2.3	2.8(2.8)	2.9(3.0)	3.4(3.2)	3.7(3.7)	4.5(5.5)
	X_1	18	2.1	2.5	2.9(2.7)	3.0(2.8)	3.1(3.2)	3.6(3.8)	3.8(4.4)
		43	0.8	1.8	2.5(2.0)	2.9(2.4)	3.0(3.1)	3.2(4.4)	3.7(4.8)
		[d] 44	0.8	1.7	2.3	2.8	3.0	3.2	3.4
Perpendicular	X_2	19	1.2	2.6	2.8(2.6)	2.9(3.0)	3.1(3.1)	3.2(3.6)	3.8(4.4)
	L	14	4.5	5.1	5.4(5.1)	5.5(5.5)	5.7(5.9)	6.1(6.8)	6.4(7.8)
	M	17	4.0	5.1	5.5(5.3)	5.6(5.5)	6.0(6.2)	6.0(6.8)	6.6(7.7)
Do	U_1	15	4.7	5.3	5.6(5.3)	5.9(5.4)	6.1(6.3)	6.5(6.9)	6.8(7.7)
	U_2	16	4.7	5.0	5.3(5.2)	5.6(5.4)	5.8(6.0)	6.0(6.5)	6.3(7.5)
Grade B (extruded)									
Parallel	A_1	28	1.8	2.0	2.1(2.0)	2.5(2.4)	2.8(2.9)	3.2(3.5)	3.4(5.1)
		40	1.7	2.1	2.4(2.1)	2.6(2.4)	2.8(2.9)	2.9(3.8)	3.0(5.7)
	A_2	33	2.1	2.4	2.6(2.4)	2.9(2.7)	3.1(3.5)	3.3(4.3)	4.1(5.7)
Perpendicular	L_1	34	3.4	3.6	4.0(3.8)	4.2(4.1)	4.6(4.7)	4.8(5.8)	5.3(7.1)
	L_2	35	3.5	3.7	4.2(3.3)	4.5(4.2)	4.8(5.2)	5.1(6.4)	5.9(7.9)
Do	U	36	4.5	4.5	4.5(4.3)	4.9(5.3)	5.3(6.1)	5.7(7.0)	6.0(8.2)
	T	37	4.1	4.5	4.6(4.0)	4.9(4.7)	5.2(5.7)	5.6(7.1)	6.4(8.0)
Grade C (molded)									
Parallel	A_1	22	5.0	5.2	5.3(5.3)	5.8(5.6)	6.0(5.9)	6.1(6.3)	6.5(7.9)
	A_2	23	4.0	4.3	4.3(4.3)	4.9(4.9)	5.0(5.1)	5.0(5.9)	[f] 5.3(7.3)
	B_1	29	4.4	4.5	4.7(4.3)	4.9(4.7)	5.3(5.3)	5.4(6.3)	5.9(7.8)
	B_2	32	4.6	4.6	5.0(4.6)	5.4(5.1)	5.3(5.9)	5.6(6.3)	5.9(7.0)
Perpendicular	T_1	24	2.9	3.6	3.8(3.0)	4.1(3.5)	4.4(4.0)	4.5(5.0)	[f] 4.5(6.7)
	T_2	25	3.2	3.2	3.6(3.5)	3.8(3.8)	4.1(4.1)	4.4(4.9)	[f] 4.8(6.6)
Do	L_1	26	2.8	3.2	3.5(3.0)	3.9(3.9)	4.3(4.2)	4.5(4.9)	[f] 4.8(6.6)
	L_2	27	3.1	3.4	3.5(3.8)	3.7(2.9)	4.2(3.6)	4.4(4.9)	[f] 5.1(6.3)

[a] Relative positions of the specimens with respect to each other and with respect to the direction of extrusion or molding pressure can be seen in fig. 1.

[b] Protected 90% platinum-10% rhodium thermocouple was used in tests 1 through 22, after which chromel-alumel thermocouples were used. Tests 1 through 13 were trials and are not included. Other missing test numbers were either faulty tests or tests of other materials.

[c] Held at 1,370° for 45 min. before cooling.

[d] Values in parentheses were obtained during cooling.

[e] Held for at least 1 hour at each of 10 temperatures between room temperature and 1,370° C.

[f] Held at 1,390° for 45 min. before cooling.

TABLE 4. *Average coefficients of linear thermal expansion for three graphites*

Grade	Orientation [a]	Average coefficients of expansion from— Room temperature to 200° C	200° to 400° C	400° to 600° C	600° to 800° C	800° to 1,000° C	1,000° to 1,200° C
A (extruded)	Parallel	2.0×10^{-6}	2.6×10^{-6}	3.0×10^{-6}	3.0×10^{-6}	3.4×10^{-6}	3.5×10^{-6}
	Perpendicular	4.9	5.2	5.5	5.7	6.0	6.5
B (extruded)	Parallel	2.0	2.4	2.5	3.0	3.1	3.5
	Perpendicular (U, T)	4.4	4.5	5.0	5.0	5.5	6.0
	Perpendicular (L_1, L_2)	3.6	4.0	4.2	4.5	4.9	5.1
C (molded)	Parallel	4.4	5.0	5.0	5.2	5.4	5.9
	Perpendicular	3.2	3.4	3.7	4.0	4.2	4.5

[a] With respect to direction of extrusion or molding pressure.

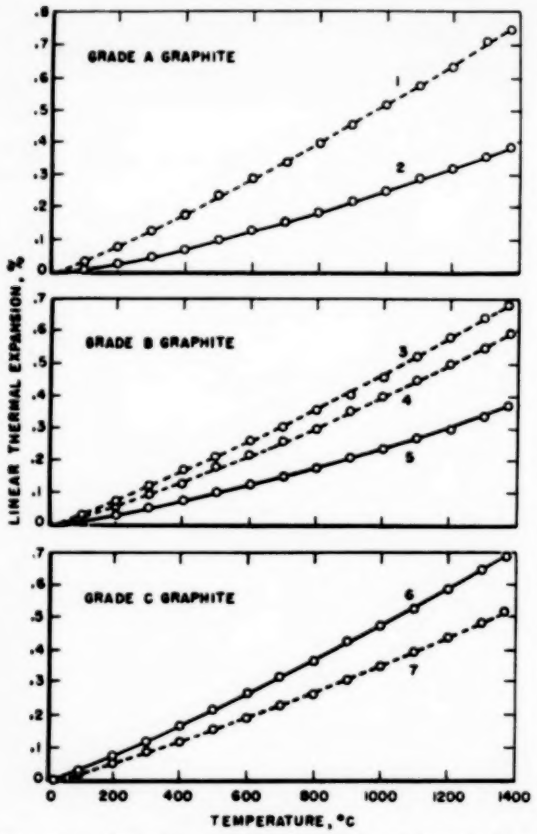

FIGURE 3. *Average linear thermal expansion of three graphites.*

Grade A graphite: 1, average of four specimens perpendicular to extrusion direction; 2, average of four specimens parallel to extrusion direction. Grade B graphite: 3, average of two specimens, C and T, perpendicular to extrusion direction; 4, average of two specimens, L_1 and L_2, perpendicular to extrusion direction; 5, average of two specimens parallel to extrusion direction; grade C graphite: 6, average of four specimens parallel to molding pressure; 7, average of four specimens perpendicular to molding pressure.

was done to evaluate the lag in temperature measurements made with the bare chromel-alumel thermocouple. The expansion measurements show that above 600° C the values during the regular test (43) were slightly larger in every case, indicating a temperature measurement lag. However, the difference does not materially affect the expansion values.

No attempt was made to compare the thermal expansion determinations of the three graphites investigated in this work with determinations of other graphites because of the wide variation of values found in the literature. The variation is attributed to the anisotropic characteristic of the crystal and the structural nature of the mass. The structure of a graphite body is unlike that of a metal or a porcelain body. The flat graphite crystals do not form impermeable solids, but rather a low density porous body. The many voids decrease the total thermal expansion considerably. The total effect would depend on the grain size and pore distribution.

The work of Nelson and Riley (see footnote 2) and Pierry (see footnote 3) indicates the degree of anisotropy of the graphite crystal. Most of the thermal expansion of the crystal is due to the expansion in the "C" direction. Along the "C" axis, the former observers (see footnote 2) found that the coefficient of expansion was about 29×10^{-6}, and along the "A" axis only 0.9×10^{-6} between 600° and 800° C.

TABLE 5. *Linear thermal expansion data obtained from repeat tests[a] of two grades of graphite*

Temperature range	Grade A specimen X_1[b]		Grade B specimen A_3	
	Test 43	Test 44	Test 28	Test 40
	Linear thermal expansion			
°C	*Percent*	*Percent*	*Percent*	*Percent*
RT[c] to 100	0.09	0.00	0.01	0.01
RT to 200	.01	.01	.03	.03
RT to 300	.03	.03	.05	.05
RT to 400	.05	.05	.07	.07
RT to 500	.07	.07	.09	.10
RT to 600	.10	.09	.11	.12
RT to 700	.13	.12	.14	.15
RT to 800	.16	.15	.16	.17
RT to 900	.19	.18	.19	.20
RT to 1,000	.22	.21	.22	.23
RT to 1,100	.25	.24	.25	.26
RT to 1,200	.28	.27	.28	.29
RT to 1,300	.32	.31	.32	.32
RT to 1,370	.35	.33	.34	.34
	Average coefficients of linear thermal expansion			
RT to 200	0.8×10^{-6}	0.8×10^{-6}	1.8×10^{-6}	1.7×10^{-6}
200 to 400	1.8	1.7	2.0	2.1
400 to 600	2.5	2.3	2.1	2.4
600 to 800	2.9	2.8	2.5	2.6
800 to 1,000	3.0	3.0	2.8	2.8
1,000 to 1,200	3.2	3.2	3.2	2.9
1,200 to 1,370	3.7	3.4	3.4	3.0

[a] Tests 43 and 44 are first and second heatings, and tests 28 and 40 are second and third heatings.

[b] Test 18 was not included because of the expected difference between expansion values due to the different method of measuring temperature.

[c] RT = room temperature.

V. Summary

1. The linear thermal expansion of specimens of three grades of graphite, cut in three directions with respect to the direction of extrusion or of molding pressure, was determined at various temperatures between 20° and 1,370° C in vacuum.

2. Specimens cut perpendicular to the extrusion direction or parallel to the direction of molding had greater expansions than specimens cut parallel to the extrusion direction or perpendicular to the molding pressure.

3. It was found that chromel-alumel thermocouples could be used successfully in vacuum at temperatures well above those at which they fail in air.

WASHINGTON, September 13, 1950.

Journal of Research of the National Bureau of Standards Vol. 47, No. 1, July 1951 Research Paper 2226

The Absolute Standardization of Radioisotopes by 4 π Counting

H. H. Seliger and L. Cavallo

The absolute counting of beta particles, using 4 π proportional flow counters, is described. This method of absolute counting has been adopted in the Radioactivity Section of the National Bureau of Standards for the primary calibration of radioisotope solution standards. Pillbox, cylindrical, and spherical 4 π flow counters have been used in the measurements, and excellent results have been obtained for each. A simple method of source mounting has been developed, and a formula is derived that determines the absorption in the source mounting regardless of knowledge of the thickness of the mounting or energy spectrum of the beta rays. Results are given of intercomparisons of P^{32} and I^{131} with other laboratories in the United States, Canada, and Great Britain that employ different types of absolute counting techniques. The applicability of the 4 π counting method to the standardization of practically all beta emitting radiosotopes is discussed.

I. Introduction

4 π proportional flow counters have been used at the National Bureau of Standards for the determination of absolute disintegration rates of radioisotopes since September 1949.[1] The principle of 2 π and 4 π counting is well known to workers in the field of radioactivity. However, several improvements have been made in the measurement technique, and since the method has not been covered previously in any detail it was felt that a description of the technique used would be of practical interest. This is particularly true in view of the excellent agreement that has been obtained with other published absolute counting methods.[2 3 4 5]

II. Technique

To eliminate the difficulties associated with Geiger counters and still retain somewhat their pulse limiting property, the 4 π counter is operated in the region of limited proportionality.[6 7] Figure 1 shows the disassembled cylindrical and pillbox 4 π counters.[8 9] The insulators are of Teflon. The stainless steel cylindrical counter was designed with the length equal to the diameter, so that the minimum path length for a particle emitted in any direction from the center of the counter would be as close as possible to the average path length. A 0.001-in. conducting diaphragm, also visible in figure 1, fits between the two halves and serves to divide the cylinder into two equal hemicylinders so that the top and bottom halves function independently. The source is mounted on the polystyrene face of a laminated film of Formvar and polystyrene [10] approximately 40 $\mu g/cm^2$ in thickness, and this thin film is cemented over the diaphragm center hole.

99 Mole percent of methane has been used as a flow gas in all counters and has been found to be preferable to an Argon-Methane flow mixture. Both the Bell Jordan linear amplifier [11] and the Brookhaven nonoverloading amplifier [12] have been used.

Two comparison tests are made on all readings taken with the 4 π counter. First integral discrimination curves are plotted for various electronic gains, keeping the counter voltage constant, and second, these discrimination curves are plotted for various counter voltages or gas gains, keeping the electronic gain constant. In both cases if the discrimination curves are flat, one can assume that all primary beta particles are being detected.

The curves shown in figures 2 and 3 illustrate the completeness of the electron collection and also the method of extrapolation to find the true disintegration rate. Even though the discrimination curves have been found in practice to be flat, it can be assumed that some primary beta particles are not detected, their number being too small to contribute appreciably to the slope of the curve. The gas amplification in this region is 10^3, and the electronic gain is 5000. The capacity of the counter circuit is approximately 10^{-11} farad. Therefore a primary beta ray must create in the neighborhood of 12 ion pairs in order for its pulse to be detected, assuming that an output pulse must be at least 10 volts in height. This would preclude from measurement all primary beta rays with energies less than 400 ev [13] which is a negligible fraction in the cases considered.

[1] Grateful acknowledgement is made to C. J. Borkowski of the ORNL, who suggested this method at a meeting of the Subcommittee on Beta and Gamma Ray Measurements and Standards of the Committee on Nuclear Science of the National Research Council in June 1949.

[2] An account of the theory and experimental application is given by L. H. Gray, British J. Radiology **22**, 677 (1949).

[3] A spherical collecting geometry is described by R. K. Clark, Columbia University, Rad. Res. Lab., N. Y. O.—1506 (March 15, 1950).

[4] A parallel plate collecting geometry is described by W. Gross and G. Failla, Phys. Rev. **79**, 209 (1950).

[5] T. B. Novey, Rev. Sci. Instr. **21**, 280 (1950).

[6] S. A. Korff, Electrons and nuclear counters (D. Van Nostrand Co., 1946).

[7] Rossi and Staub, Ionization chambers and counters (McGraw-Hill, 1949).

[8] Rossi and Staub, Ionization chambers and counters, p. 130 (McGraw-Hill, 1949), illustrate a 4 π counter used for α particles.

[9] A diagram of a 4 π counter is shown in a review article on beta counting of J. L. Putnam, United Kingdom, AERE-NR-318 (Feb. 24, 1949.)

[10] L. M. Fry and R. T. Overman, Atomic Energy Com. Doc. 1800 (Jan. 30, 1948).

[11] P. R. Bell and W. H. Jordan, Rev. Sci. Instr. **18**, 703 (1947).

[12] W. Bernstein, H. G. Brewer, Jr., and W. Rubinsin, Nucleonics **6**, 39 (1950).

[13] This agrees roughly with the calculations of R. Ballentine, private communication.

FIGURE 1. *Disassembled cylindrical and pillbox 4 π counters.*

The "O" ring Teflon insulators are clearly visible. In the foreground are the diaphragms and the 0.001 in. aluminum source mounting ring. The relative scale is given by the [centimeter rule] in the [center].

III. Source Mounting

Experiments have been conducted to determine the amount of absorption, if any, due to the film between the source and the lower half of the 4π counter. The number of particles counted by the top half of the counter connected separately will be

$$N_{\text{top}}=\frac{N_0}{2}[1+B_F+(1-\tau)B_w(b)], \tag{1}$$

where N_0 is the true disintegration rate of the source, B_F is the percentage backscattering from the film, τ is the fractional absorption in the film, and $B_w(b)$ is the percentage backscattering due to the walls in the bottom half. The number of particles counted by the bottom half connected separately will be

$$N_{\text{bottom}}=\frac{N_0}{2}[(1-\tau)+(1+B_F)B_W(t)]. \tag{2}$$

From symmetry considerations $B_W(t)=B_W(b)=B_W$. The factor B_F can be neglected when the film is thin and of low atomic number, so that (1) becomes

$$N_{\text{top}}=\frac{N_0}{2}[1+B_W-\tau B_W]. \tag{3}$$

In the bottom half, again assuming $B_F=0$, one obtains

$$N_{\text{bottom}}=\frac{N_0}{2}[(1-\tau)+B_W]. \tag{4}$$

Putting this in the form $y=ax+b$ gives

$$N_{\text{bottom}}=-\frac{N_0}{2}\tau+\frac{N_0}{2}(1+B_W). \tag{5}$$

With the thin films under consideration it can be assumed that the absorption is directly proportional to the film thickness. Equation (5) can be used to determine the absorption correction graphically. A more direct method of determining τ can be deduced from eq (3) and (4).

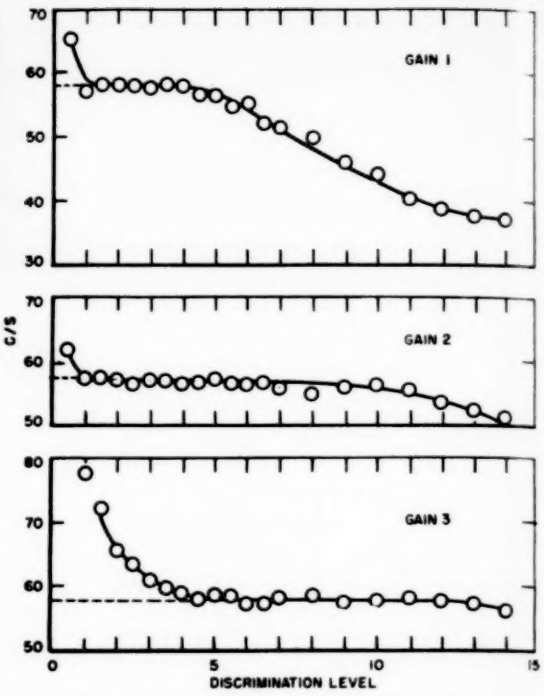

FIGURE 2. *Counting rate as a function of discrimination level for increasing steps in amplifier gain.*

The source is P^{32} in the pillbox 4 π counter. The voltage on the counter was held constant during all runs. Gain 2 is twice that of gain 1, and gain 3 is 4 times that of gain 1.

$$N_{\text{top}}-N_{\text{bottom}}=\frac{N_0}{2}\tau(1-B_W). \tag{6}$$

The actual counting rate observed with top and bottom halves connected together is

$$N_{tb}=N_0(1-\tau/2) \tag{7}$$

so that (6) becomes, if one lets $N_{\text{top}}-N_{\text{bottom}}=\Delta$

$$\tau=\frac{N_{\text{top}}-N_{\text{bottom}}}{N_{tb}-N_{\text{bottom}}}=\frac{\Delta}{N_{tb}-N_b}, \tag{8}$$

and similarly

$$B_W=\frac{N_{\text{top}}+N_{\text{bottom}}}{N_{tb}}-1. \tag{9}$$

The absolute counting rate is then obtained by substituting (8) into

$$N_0=\frac{N_{tb}}{1-\tau/2}. \tag{10}$$

Thus by taking three different readings of the same source on a single film it is possible to determine the absorption by the film. This proves extremely useful for low energy beta particles.

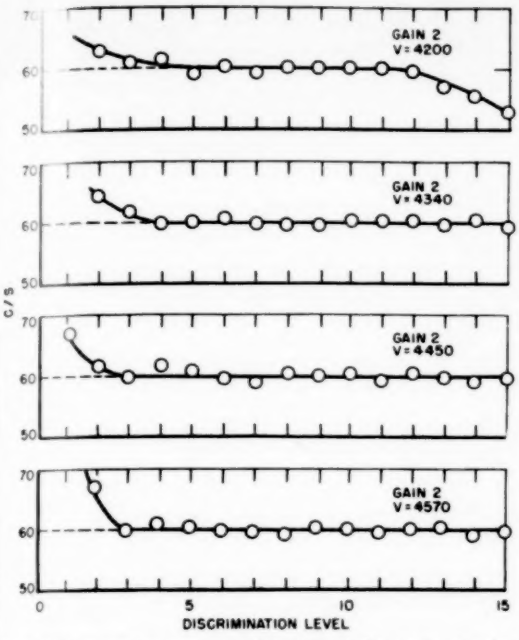

FIGURE 3. *Counting rate as a function of discrimination level for increasing steps in voltage.*

The source is I^{131} in the cylindrical 4 π counter. The electronic gain was held constant, and discrimination curves were run for various voltages on the counter.

Experiments have been performed with P^{32}, I^{131}, and CO^{60} using film thickness in multiples of approximately 40 $\mu g/cm^2$, and excellent agreement has been obtained between eq (5) and (9). As an example of the corrections necessary for absorption, for a film of 40 $\mu g/cm^2$ thickness $\tau = 0$ for P^{32}, $\tau = 0.01$ for I^{131}, and $\tau = 0.05$ for Co^{60}. In the cases of I^{131}, Co^{60}, and C^{14} polystyrene films as low as 10 $\mu g/cm^2$ in thickness have been used. More complete data are now being taken for Co^{60}, C^{14}, and Na^{22}.

Previously when thin films were used as source mounts it was thought necessary to sputter a thin conducting layer of gold on both sides of the film so as not to disburb the collecting fields. However, if the area of the dielectric film is made small compared to the total area of the conducting diaphragm these collecting fields will be perturbed only slightly with no loss in efficiency. All results shown for the cylinder and spherical counters have been obtained using nonconducting plastic films, and table 1 shows the agreement for P^{32} between the cylinder counter in which such thin films were used as source mountings and the pillbox counter in which the sources were mounted on a conducting film of aluminum.

IV. Source Thickness

Equation (10) previously derived is valid, provided there is no self-absorption in the source itself. In any case the effect of self-absorption in a source of the same thickness as the mounting film is not as serious as the absorption in the film. The source is an evaporation covering a small area, and therefore there is a slowing down of electrons rather than complete stopping, as in the case of the film where some of the emitted electrons can have a long path length. Since the specific ionization curve has a maximum at approximately 7,000 ev, this slowing down usually adds to the total ionization. The effect of self-absorption will be more important in the cases of Co^{60}, C^{14}, and S^{35}.

V. Results and Conclusions

P^{32} and I^{131} sources have been calibrated in all three types of 4 π counters. The internal consistency of repeated measurements in all cases is within 1 percent. The results of some determinations made using the three types of counters are shown in table 1.

It is felt that the cylindrical 4π counter has slightly better collecting fields than either the sphere or the pillbox. This does not affect in any way the absolute counting of the beta particles, but is evidenced by a slightly longer voltage plateau for the cylinder. However, the spherical counter is by far the most economical, the total cost including modifications and mounting stand being approximately $80.00.[14]

The cylindrical 4π counter is at present being used for all standardizations.

The results of a series of intercomparisons of P^{32} and I^{131} solutions originally calibrated at the National Bureau of Standards and distributed to other laboratories are given in table 2.

The technique described above is applicable to any beta emitting radioisotope, regardless of the complexity of the disintegration scheme. Isomeric states or gaseous isotopes may cause difficulties. However all effects due to gamma-ray line spectra, gamma-ray conversion electrons, or electrons produced in the gas or the walls of the counter by gamma rays only add to the total ionization per disintegration and are counted as one ionization pulse. A high specific activity source is not required, since such low counting rates can be recorded that water dilution is practically always possible. Most important, the method is self-consistent. Any improper mounting or positioning of the source is shown up immediately by a large value for τ. Any other malfunctioning will not result in flat discrimination curves. There are no further meter or condenser calibrations necessary. Equation (10) gives the true disintegration rate regardless of changes in temperature, humidity, pressure inside the chamber, or voltage.

The 4 π counter method is at present being used to calibrate sources of C^{14}, Na^{22}, and Co^{60}. As soon as practicable S^{35}, Na^{24}, Au^{198}, and Sr^{90} will be calibrated.

[14] Hemispheres are available from Nuclear Measurements Corp., 3339 Central Ave., Indianapolis, Ind.

TABLE 1. *Results of comparison measurements made of P^{32} and I^{131} in all three types of 4π counters*

[The errors indicated are statistical standard deviations]

Isotope and date of distribution	Counting rate	Type of counter	Type of amplifier	Result
	Counts/sec			*dps/ml*$\times 10^{-3}$
I^{131} 6/8/50	Less than 100	Pillbox	Bell Jordan	300±2
	do	Cylinder	do	298±1
	do	do	do	239±0.63
P^{32} 9/14/50	do	do	Brookhaven	241±0.40
	do	Sphere	Bell Jordan	240±0.24
	Approximately 900	Cylinder	Brookhaven	239±1.0

TABLE 2. *Results of independent intercomparisons of P^{32} and I^{131} made by other researchers using absolute counting techniques*

The errors, except for Dr. Hawkings' figures, are estimated inaccuracies in absolute value

Isotope and date of distribution	Activity dps/ml$\times 10^{-3}$				
	NBS 4π counter	L. H. Gray ionization chamber	R. C. Hawkins 4π counter [a]	Oak Ridge defined solid angle	G. Failla parallel plate chamber
I^{131} 12/13/49	250±7.5	248±6.2	.....	.....	.....
P^{32} 2/27/50	327±9.8	.....	322±1.4	317	.....
I^{131} 6/8/50	299±9.0	300±1.2	304±1.7	.....	.....
P^{32} 9/14/50	240±7.0	229±5	243±1	235±5	241±7

[a] R. C. Hawkings, Can. Nat. Research Coun., Chalk River, Canada, uses a spherical 4π counter similar to the NBS counter for his determinations.

The authors are indebted to R. C. Hawkings for his cooperation in the intercomparisons of the P^{32} and I^{131} solution standards, and to L. H. Gray, Hammersmith Hospital, London, Eng., G. Failla, Columbia University School of Medicine, E. I. Wyatt and J. A. Cox of the Oak Ridge National Laboratory for their interest and cooperation. Thanks is also given to L. Costrell of the Nucleonics Instrumentation Section for the construction of the electronic equipment, for his valuable assistance in setting up the pillbox 4π counter in September of 1949, and for his continued interest.

WASHINGTON, December 18, 1950.

Journal of Research of the National Bureau of Standards Vol. 47, No. 1, July 1951 Research Paper 2227

A Method of Gradients for the Calculation of the Characteristic Roots and Vectors of a Real Symmetric Matrix[1]

Magnus R. Hestenes and William Karush

Let A be a real symmetric matrix, $\mu(x)$. The Rayleigh quotient formed with a vector x, and $\xi(x)$ the gradient vector of $\mu(x)$. The method of gradients consists in an infinite iteration of the operation $\bar{x}=x-\alpha(x)\,\xi(x)$. The convergence of the procedure is proved for several choices for $\alpha(x)$, and the rate of convergence is studied extensively for one particular $\alpha(x)$. The directions of the vectors in the sequence are seen to tend to that of the characteristic vector belonging to the lowest characteristic value. The method can be used for a numerical determination of all characteristic vectors and values.

I. Introduction

With a real symmetric matrix

$$A=(a_{ij}),\ a_{ij}=a_{ji},\qquad i,j=1,2,\cdots,n,$$

is associated the Rayleigh quotient

$$\mu(x)=\frac{(Ax,x)}{(x,x)},\qquad x\neq 0,$$

whose critical points y are the characteristic vectors of A. The gradient of μ has the direction of [2]

$$\xi(x)=Ax-\mu x. \tag{1}$$

Suppose now, given a vector x, we wish to modify x to obtain a better approximation $\bar{x}$ to a characteristic vector $y_{\min}$ corresponding to the minimum characteristic root $\lambda_{\min}=\min\ \mu(x)$. It is natural to form

$$\bar{x}=x-\alpha\xi,\qquad \alpha>0, \tag{2}$$

where α may depend upon x. Similarly, to approximate to a characteristic vector $y_{\max}$ corresponding to $\lambda_{\max}=\max\ \mu(x)$ we form

$$x=x+\alpha\xi,\qquad \alpha>0.$$

In the present paper we describe several convergent iterative methods based upon this gradient process, and investigate the convergence to the characteristic roots and vectors of A. The results apply to an arbitrary, real symmetric matrix. The methods can be phrased to yield directly $y_{\min}$ or $y_{\max}$, as one wishes. For convenience we direct our attention mainly to $y_{\min}$.

In prescribing a gradient method one must specify how the number α is to be chosen at each stage x^i of the iteration. It is shown that the vectors x^i converge to the appropriate characteristic vector if α is any positive constant (independent of i) less than $2/M$, $M=\lambda_{\max}-\lambda_{\min}$. The bulk of the theory is concerned with this "method of fixed α." If we impose the stricter requirement $\alpha<1/M$, we obtain in addition that the gradients ξ^i converge in direction (i. e., the unit vectors $\xi^i/|\xi^i|$ converge) to a characteristic vector; in fact the method will be generalized to yield all of the characteristic vectors of A. As would be expected, the nature of the convergence is essentially geometric.

A second "method of optimum α" is treated in which α, which now depends upon x, is selected in a certain "best" way. In this method the approximations x^i converge to a characteristic vector, but the gradients ξ^i fail to converge in direction.[3]

The well-known method [4] of forming powers A^ix can be interpreted as a gradient method in which α is chosen as $-1/\mu$. Here convergence, in general, is to a characteristic vector corresponding to the characteristic root of maximum absolute value. We remark that, commonly, the gradients ξ^i converge in direction to a characteristic vector corresponding to a root of next highest absolute value.

The chief virtue of the gradient methods seems to lie in their simplicity. They are not put forth as rapid procedures for a hand computer, but rather as processes that might be adapted to automatic computing machinery. A survey of methods for calculating characteristic roots and vectors of (more general) matrices may be found in Hammersley.[5]

II. Properties of Symmetric Matrices

In this section we collect for reference some well-known facts on symmetric matrices.

[1] The preparation of this paper was sponsored (in part) by the Office of Naval Research.

[2] For convenience we shall refer to ξ as the gradient.

[3] An extension of the method of optimum α to more general problems has been outlined by L. V. Kantorovitch, On an effective method of solving extremal problems for quadratic functionals, C. R. (Doklady) Acad. Sci., URSS (N. S.), **48**, 455–460 (1945). These results are closely related to some unpublished work of M. R. Hestenes.

[4] See H. Hotelling, Analysis of a complex of statistical variables into principal components, J. Educ. Psych. **24**, 417–441 and 498–520 (1933). In this paper, Hotelling treats the symmetric matrix. For the extension to nonsymmetric matrices, see A. C. Aitkin, Studies in practical mathematics II. The evaluation of the latent roots and latent vectors of a matrix, Proc. Roy. Soc. Edinburgh [A] **57**, 269–304 (1937).

[5] The numerical reduction of non-singular matrix pencils, Phil. Mag. **40**, 783–807 (1949).

Consider, for the moment, the space of complex vectors $x=(b_1, b_2, \ldots, b_n)$ over the field of complex numbers. A number Λ is called a characteristic root (number, value) of an arbitrary matrix B in case there exists a nonnull vector x such that

$$Bx=\Lambda x.$$

The vector x is called a characteristic vector; we shall say that it belongs to the characteristic root Λ.

For a real symmetric matrix A the characteristic roots are real and the characteristic vectors can be chosen to be real. Accordingly we henceforth limit ourselves to the field of real numbers and to real vectors. We let $\mathscr{A}$ denote the space of real vectors x. We use "(x, y)" for inner product, and "$|x|$" for length in this space. We note a fundamental property of A:

$$(Ax, y)=(x, Ay) \tag{3}$$

Let the distinct characteristic roots of A be denoted by

$$\Lambda_1<\Lambda_2<\ldots<\Lambda_\rho, \quad \rho\geq 2,$$

(eliminating the trivial case $\rho=1$). With each Λ_k is associated the linear subspace $\mathscr{M}_k$ of $\mathscr{A}$, which is the set of all characteristic vectors belonging to Λ_k, together with the null vector. The dimension of $\mathscr{M}_k$ is the order of Λ_k. The subspaces $\mathscr{M}_k$ are mutually orthogonal and span the space $\mathscr{A}$.

To each nonnull vector x we attach the number

$$\mu(x)=\frac{(Ax, x)}{(x,x)}.$$

The function μ is homogeneous of degree zero;

$$\mu(bx)=\mu(x), \quad b\neq 0.$$

For a characteristic vector y belonging to a characteristic root Λ,

$$\mu(y)=\Lambda.$$

We have the following well-known relations between μ and the characteristic roots.

$$\Lambda_1=\min_x \mu(x), \quad \Lambda_\rho=\max_x \mu(x), \qquad x\neq 0$$

or, equivalently,

$$\Lambda_1=\min_x \mu(x), \quad \Lambda_\rho=\max_x \mu(x), \qquad |x|=1.$$

More generally,

$\Lambda_k=\min_x \mu(x)$, $x\neq 0$ and orthogonal to $\mathscr{M}_1, \ldots, \mathscr{M}_{k-1}$;

$\Lambda_k=\max_x \mu(x)$, $x\neq 0$ and orthogonal to $\mathscr{M}_{k+1}, \ldots, \mathscr{M}_\rho$.

To study the behavior of subspaces of $\mathscr{A}$ under the matrix A it is convenient to think of A as a linear operator, without regard for coordinate representation. In general, if B is a symmetric linear operator on a finite dimensional space $\mathscr{B}$, then the statements of the preceding three paragraphs hold (by "symmetry" is meant the property of (3)). Now suppose that $\mathscr{B}$ is an m-dimensional subspace of $\mathscr{A}$ and let B be the operator A with domain restricted to $\mathscr{B}$. If $\mathscr{B}$ is invariant under A, i. e., if x belonging to $\mathscr{B}$ implies Ax belongs to $\mathscr{B}$, then B is a symmetric linear operator on $\mathscr{B}$. As such, B has m characteristic roots (counting multiplicities), and the roots and vectors of B are roots and vectors of A.

The characteristic roots of the matrix A are the solutions of the polynomial equation

$$|\lambda I-A|=0.$$

The multiplicity of a solution Λ_k is precisely the dimension of the subspace $\mathscr{M}_k$, i. e., the order of Λ_k. We remark that if A is regarded as a linear operator, then this equation, formed with any matrix representation of A, yields the characteristic roots.

In our work we shall be dealing always with a fixed initial vector x^0. Let x^0 have nonnull projections in $\mathscr{M}_{k_1}, \ldots, \mathscr{M}_{k_r}$ and only in these subspaces $\mathscr{M}_k$. We may represent x^0 uniquely in the form

$$x^0=a_1^0 y_1+\cdots+a_r^0 y_r \tag{4}$$

$$a_j^0>0, \quad |y_j|=1, \quad y_j\in \mathscr{M}_{k_j}, \quad j=1, 2, \cdots, r\geq 2.$$

(We exclude the trivial case $r=1$.) Let $\mathscr{A}_r$ be the r dimensional subspace spanned by $y_1, \ldots, y_r$;

$$\mathscr{A}_r=(y_1, y_2, \cdots, y_r).$$

Then $\mathscr{A}_r$ is an invariant subspace under A with orthonormal basis $(y_1, y_2, \ldots, y_r)$. The characteristic roots of A relative to this invariant subspace (i. e. the roots of the linear operator A with domain restricted to $\mathscr{A}_r$) are

$$\lambda_1<\lambda_2<\cdots<\lambda_r$$

where

$$\lambda_j=\Lambda_{k_j}, \qquad j=1, \cdots, r.$$

In the subspace each root λ_j has order one, and has as corresponding characteristic vector, y_j. In the following pages, after having selected an initial vector x^0, and thus determined the invariant subspace $\mathscr{A}_r$, we shall be dealing exclusively with vectors in this subspace.

III. Determination of α

This section is discursive; the theorems of the following sections do not logically depend upon it.

A direct calculation shows that the gradient of μ is

$$\frac{2}{(x, x)}[Ax-\mu(x)x],$$

hence has the direction of ξ, given by eq (1). As remarked, we shall refer to ξ as the gradient. Observe the relations

$$(x, \xi)=0, \quad (Ax, \xi)=(\xi, \xi). \tag{5}$$

The gradient is the direction in which μ locally increases most rapidly. Thus if we form $\overline{x}$ from x as in eq (2), with $\alpha>0$ sufficiently small, we should expect $\mu(\overline{x})$ to approximate λ_1 more closely than $\mu(x)$. Beginning now with $\overline{x}$ and, say, $\overline{\alpha}$, we obtain from eq (2) a next approximation $\overline{\overline{x}}$, etc. The question then is to specify α systematically.

By direct computation we find

$$\mu(x)-\mu(\overline{x})=\frac{\alpha(2-\alpha s)}{1+\alpha^2 t^2}\, t^2 \tag{6}$$

where

$$s=\mu(\xi)-\mu(x), \quad t=\frac{|\xi|}{|x|}.$$

That is,

$$\mu(x)-\mu(\overline{x})=f(\alpha)t^2, \quad f(\alpha)\equiv f(\alpha, x)=\frac{\alpha(2-\alpha s)}{1+\alpha^2 t^2}. \tag{7}$$

A natural requirement on $\alpha(x)$ in order to expect convergence of μ to λ_1 is then

$$f(\alpha(x))>0.$$

A possible choice is $\alpha=$ const. with

$$0<\alpha<\frac{2}{M}, \tag{8}$$

where

$$M=\Lambda_\rho-\Lambda_1$$

is the spread of the characteristic roots of A. Since $s\leq M$ it is easily seen that $f(\alpha)>0$. The method of fixed α, stemming from this observation, is treated in sections 5 and 6 and generalized in sections 10 and 11.

Another possibility for α is that number $\gamma=\gamma(x)$ which maximizes $f(\alpha)$. From $f'(\gamma)=0$ we obtain

$$t^2\gamma^2+s\gamma-1=0, \tag{9}$$

from which

$$\gamma=\frac{2}{s+\sqrt{s^2+4t^2}}>0. \tag{10}$$

Computation shows that $f(\gamma)=\gamma$. Hence

$$\mu(x)-\mu(\overline{x})=\gamma t^2, \quad \overline{x}=x-\gamma\xi.$$

The method of optimum α, based on this approach, is treated in section 7. We note the following formula:

$$\gamma=\frac{1}{\mu(\xi)-\mu(\overline{x})}. \tag{11}$$

To verify this we write eq (10) as $\gamma(\gamma t^2+s)=1$, and note that the quantity in parentheses equals $\mu(\xi)-\mu(x)$.

Remarks on the power method, i. e., forming successive iterates $A^i x$, will be made in section 8. In that section and in the later section 12 modifications of the methods of fixed and optimum α will be suggested.

Before proceeding to any specific gradient method we formulate some general results in the next section.

IV. General Gradient Method

We suppose given a real-valued function γ such that $\gamma(x)$ is defined whenever $x\neq 0$ and x is not a characteristic vector, i. e., whenever $x\neq 0$ and $\xi\neq 0$. We require that

$$f(\gamma(x))>0.$$

(See, e. g., the particular function γ determined by eq (10). Beginning with a vector x^0, expressed in the form (4), we construct the sequence of vectors $\{x^i\}$, $i=0, 1, 2, \ldots$, according to

$$x^{i+1}=x^i-\gamma^i\xi^i, \text{ where } \xi^i=\xi(x^i), \ \gamma^i=\gamma(x^i). \tag{12}$$

This is equivalent to

$$x^{i+1}=(1+\gamma^i\mu^i)x^i-\gamma^i Ax^i, \text{ where } \mu^i=\mu(x^i). \tag{13}$$

We must assure ourselves that $\{x^i\}$ is well defined. We note first that if $x^i\neq 0$ and $\xi^i\neq 0$, then x^{i+1} is defined and not null. For, under the hypotheses, x^{i+1} is defined and by eq (12),

$$|x^{i+1}|^2=|x^i|^2+\gamma^{i2}|\xi^i|^2.$$

We thereby distinguish two types of sequences. The first, a trivial case, is one such that for some first integer k,

$$\xi^k=0.$$

We terminate the induction at $i=k$; the sequence $\{x^i\}$, $i=0, 1, \cdots, k$, is finite with the last vector x^k a characteristic vector. The second type is characterized as a sequence that is not of the first type. In this case $\{x^i\}$ and $\{\xi^i\}$ are well-defined infinite sequences and

$$\xi^i\neq 0, \qquad i=0, 1, 2, \cdots.$$

To enable us to formulate statements that hold simultaneously for finite and infinite sequences we make the following agreements: (1) if $\{p^i\}$ is an arbitrary finite sequence, then $\lim_{i\to\infty} p^i$ denotes the last member of the sequence; (2) any statement or formula involving an index is to be limited to meaningful values of the index.

From eq (7),

$$\mu^i-\mu^{i+1}=f(\gamma^i)t^{i2}>0, \quad t^i=\frac{|\xi^i|}{|x^i|}. \tag{14}$$

Since μ^i is bounded from below, it follows that there is a number ν such that

$$\lim_{i\to\infty} \mu^i = \nu. \tag{15}$$

The vectors x^i and ξ^i lie within the invariant subspace $\mathscr{M}=(y_1, y_2, \cdots, y_r)$; their expansions in terms of the characteristic vectors y_j are readily determined from eq (1) and (12). We find

$$\left.\begin{aligned} x^i &= a_1^i y_1 + a_2^i y_2 + \cdots + a_r^i y_r, \\ a_1^i &= [1+\gamma^{i-1}(\mu^{i-1}-\lambda_1)]\, a_1^{i-1} \\ a_j^i &= [1+\gamma^{i-1}(\mu^{i-1}-\lambda_j)]\, a_j^{i-1}, \qquad j=1,2,\cdots,r, \end{aligned}\right\} \tag{16}$$

and

$$\xi^i = (\lambda_1-\mu^i)a_1^i y_1 + \cdots + (\lambda_r-\mu^i)a_r^i y_r. \tag{17}$$

Thus

$$\left.\begin{aligned} |x^i|^2 &= a_1^{i\,2} + a_2^{i\,2} + \cdots + a_r^{i\,2}, \\ |\xi^i|^2 &= (\lambda_1-\mu^i)^2 a_1^{i\,2} + \cdots + (\lambda_r-\mu^i)^2 a_r^{i\,2}. \end{aligned}\right\} \tag{18}$$

Theorem 4.1. Suppose γ^i is such that

$$0<K_1\le f(\gamma^i) \text{ and } 0<\gamma^i,$$

for some constant K_1. Then

$$\lim_{i\to\infty} \frac{x^i}{|x^i|} = y_1 \text{ and } \lim_{i\to\infty} \mu^i = \lambda_1.$$

We limit ourselves to the case when the sequence $\{x^i\}$ is infinite; by appropriate simplification the proof below applies as well to the finite sequence.

From eq (14)

$$t^{i2} \le \frac{1}{K_1}(\mu^i - \mu^{i+1}). \tag{19}$$

Since $\mu^i \to \nu$, it follows that

$$t^i \to 0.$$

Let

$$b_j^i = \frac{|a_j^i|}{|x^i|}.$$

Then by eq (18)

$$t^{i2} = \frac{|\xi^i|^2}{|x^i|^2} = \sum_{j=1}^{r} (\lambda_j-\mu^i)^2 b_j^{i\,2}, \quad \sum_{j=1}^{r} b_j^{i\,2} = 1. \tag{20}$$

Consequently

$$\lim_{i\to\infty} (\lambda_j-\mu^i)^2 b_j^{i\,2} = 0, \qquad j=1,2,\cdots,r.$$

Since $(\lambda_j-\mu^i)\to(\lambda_j-\nu)$ it follows from the last equation of (20) that for some specific value l of the index j,

$$\mu^i \to \lambda_l, \quad b_l^i \to 1, \quad b_j^i \to 0 \text{ for } j\ne l.$$

We wish to show that $l=1$. Suppose $l>1$. Then

$$\frac{b_1^i}{b_l^i} = \frac{|a_1^i|}{|a_l^i|} \to 0.$$

From the monotonicity of μ^i, we have $\mu^i>\lambda_l>\lambda_1$. From eq (16)

$$\frac{|a_1^{i+1}|}{|a_l^{i+1}|} = \frac{|1+\gamma^i(\mu^i-\lambda_1)|}{|1+\gamma^i(\mu^i-\lambda_l)|} \cdot \frac{|a_1^i|}{|a_l^i|} \ge \frac{|a_1^i|}{|a_l^i|}.$$

Hence

$$\frac{|a_1^i|}{|a_l^i|} \ge \frac{|a_1^0|}{|a_l^0|} > 0 \text{ for } i=0,1,2,\cdots.$$

This contradiction shows that $l=1$. We now have

$$\mu^i \to \lambda_1, \quad \frac{a_1^i}{|x^i|} \to 1, \quad \frac{a_j^i}{|x^i|} \to 0, \quad j\ne 1,$$

since

$$a_1^i \ge a_1^0 > 0.$$

The theorem now follows from the first equation of (16).

Theorem 4.2. If, in addition to the hypotheses of theorem 4.1, the sequence γ^i satisfies the condition

$$\gamma^i \le K_2, \quad K_2 = \text{const.},$$

then the conclusion of that theorem may be strengthened to

$$\lim_{i\to\infty} x^i = L y_1,$$

where L is a positive constant.

By theorem 4.1 it is sufficient to show that $|x^i|\to L$. From eq (12),

$$|x^{i+1}|^2 = |x^0|^2[1+(\gamma^0 t^0)^2]\ [1+(\gamma^1 t^1)^2]\ \cdots\ [1+(\gamma^i t^i)^2].$$

It is well known that the product on the right converges if the series $\sum_0^\infty (\gamma^i t^i)^2$ does. From eq (19) the series Σt^{i2} converges, and hence the first series does, by the boundedness of γ^i. This completes the proof.

V. The Method of Fixed α

In this gradient method we choose $\gamma(x)$ to be a constant α satisfying eq (8); thus,

$$\alpha = \frac{\beta}{M}, \qquad 0<\beta<2. \tag{21}$$

It follows that

$$f(\alpha) \ge \frac{\alpha(2-\beta)}{1+\alpha^2 t^2}, \quad x\ne 0.$$

From eq (1) we derive

$$|\xi| \leq 2K|x|,$$

where K is the bound of the matrix A. Hence t is bounded from above. Thus $f(\alpha)$ is bounded from below by a positive constant. Theorem 4.2 is now applicable.

Theorem 5.1. Let γ^i be chosen as a constant α satisfying eq (21). Then

$$\lim_{i\to\infty} x^i = L y_1, \qquad \lim_{i\to\infty} \mu^i = \lambda_1.$$

For later use we note that

$$\frac{|x^i|}{a_1^i} \to 1, \tag{22}$$

which was established in the proof of theorem 4.1.

We propose to show now that under a strengthening of condition (21) the gradients ξ^i converge in direction to the second characteristic vector y_2. The new condition is

$$\alpha = \frac{\beta}{M}, \qquad 0 < \beta < 1. \tag{23}$$

Lemma 5.1. Under the condition (23) the sequence $\{x^i\}$ is infinite and

$$a_j^i > 0, \quad \xi^i \neq 0, \quad j = 1, 2, \ldots, r; \; i = 0, 1, 2, \ldots.$$

From eq (16)

$$a_j^i = \{1 + \alpha(\mu^{i-1} - \lambda_j)\} a_j^{i-1}. \tag{24}$$

By eq (23), the expression in braces is positive. Hence, from $a_j^0 > 0$, we have $a_j^i > 0$. From the second equation (18) and the assumption $r \geq 2$, it follows that $\xi^i \neq 0$. The last inequality assures that $\{x^i\}$ is infinite.

We introduce the following notation

$$\delta_j = 1 - \alpha(\lambda_j - \lambda_1) = 1 - \beta \frac{\lambda_j - \lambda_1}{M}, \quad j = 1, 2, \ldots, r. \tag{25}$$

Thus, under the condition (23),

$$0 < \delta_r < \delta_{r-1} < \ldots < \delta_2 < \delta_1 = 1.$$

Lemma 5.2 Under the condition (23),

$$\lim_{i\to\infty} \frac{a_j^{i+1}}{a_j^i} = \delta_j,$$

$$\lim_{i\to\infty} \frac{a_k^i}{a_j^i} = 0,$$

for $j < k$.

The first equation follows from (24) and the fact that $\mu^i \to \lambda_1$. To prove the second equation notice that for $j < k$,

$$\frac{a_k^{i+1}}{a_j^{i+1}} \div \frac{a_k^i}{a_j^i} \to \frac{\delta_k}{\delta_j} \text{ and } 0 < \frac{\delta_k}{\delta_j} < 1.$$

Hence, for any $\epsilon > 0$ there is a constant K and index i_0 such that

$$\frac{a_k^i}{a_j^i} < K\left(\frac{\delta_k}{\delta_j} + \epsilon\right)^i, \qquad i \geq i_0.$$

Choose ϵ so that the number in parentheses is less than 1. This completes the proof.

Lemma 5.3 Under the condition (23),

$$\lim_{i\to\infty} \frac{(\mu^i - \lambda_1) a_1^{i\,2}}{a_1^{i\,2}} = (\lambda_2 - \lambda_1), \quad \lim_{i\to\infty} \frac{|\xi^i|}{a_2^i} = \lambda_2 - \lambda_1.$$

From the orthogonality relation $(x^i, \xi^i) = 0$ and eq (16) and (17) we obtain

$$(\lambda_1 - \mu^i) a_1^{i\,2} + (\lambda_2 - \mu^i) a_2^{i\,2} + \ldots + (\lambda_r - \mu^i) a_r^{i\,2} = 0.$$

Divide by $a_1^{i\,2}$ and take the limit as $i \to \infty$. The first conclusion of the lemma then follows from lemma 5.2. Using this result we establish the second conclusion in a similar way by means of eq (18).

Theorem 5.2. Let the constant α be chosen to satisfy (23). Then

$$\lim_{i\to\infty} \frac{\xi^i}{|\xi^i|} = y_2, \qquad \lim_{i\to\infty} \mu(\xi^i) = \lambda_2.$$

Divide both sides of eq (17) by a_2^i. Using lemmas 5.2 and 5.3 we obtain

$$\lim_{i\to\infty} \frac{\xi^i}{a_2^i} = (\lambda_2 - \lambda_1) y_2.$$

It follows that

$$\frac{\xi^i}{|\xi^i|} = \frac{\xi^i}{a_2^i} \cdot \frac{a_2}{|\xi^i|} \to y_2,$$

as desired. Also,

$$\mu(\xi_1) = \mu\left(\frac{\xi^i}{|\xi^i|}\right) \to \mu(y_2) = \lambda_2.$$

This completes the proof.

We remark that for the proofs of theorems 5.1 and 5.2 the conditions (21) and (23) could have been relaxed by using $\lambda_r - \lambda_1$ in place of $M = \Lambda_\rho - \Lambda_1$.

We conclude this section with a reformulation of theorems 5.1 and 5.2 for convergence to the two highest characteristic vectors y_r, y_{r-1}.

Theorem 5.3. Let the constant α satisfy eq (21). Define the sequence $\{x^i\}$ by the recursion formula

$$x^{i+1} = x^i + \alpha \xi^i.$$

Then

$$\lim_{i\to\infty} x^i = L y_r, \quad \lim_{i\to\infty} \mu(x^i) = \lambda_r, \quad L>0.$$

If, further, α satisfies eq (23,) then

$$\lim_{i\to\infty} \frac{\xi^i}{|\xi^i|} = -y_{r-1}, \quad \lim_{i\to\infty} \mu(\xi^i) = \lambda_{r-1}.$$

This result is obtained by replacing A by $B=-A$ in theorems 5.1 and 5.2, and noticing that the characteristic roots become

$$-\lambda_r < -\lambda_{r-1} < \cdots < -\lambda_1$$

with corresponding characteristic vectors y_r, y_{r-1}, $\ldots$, y_1.

VI. Rate of Convergence for Fixed α

In this section we investigate the rate of convergence of the sequences $\{x^i\}$, $\{\xi^i\}$ and related sequences; in the rest of this section we assume eq (23) holds. For convenience we introduce the so-called "ratio" of a sequence of real numbers as a measure of speed of convergence to 0, and develop some elementary properties of sequences with ratios. This notion will also be useful in our later generalization of the method of fixed α.

Let us agree that a sequence $\{b^i\}$ of real numbers will be called positive (negative) if $b^i>0 (<0)$ for $i\geq i_0$, i_0 fixed. We shall understand that the sequence is monotonic in case it is monotonic for $i\geq i_0$.

Definition. A sequence $\{b^i\}$ of real numbers will be said to have the ratio κ in case

$$\lim_{i\to\infty} \frac{b^{i+1}}{b^i} = \kappa \text{ with } \kappa>0.$$

Necessarily a sequence with a ratio is either positive or negative.

The reason for this definition lies in the next lemma, which is essentially a rephrasing of Lemma 5.2.

Lemma 6.1. The sequences $\{a_j^i\}$, $\{a_k^i/a_j^i\}$ $(j,k=1, \cdots, r)$ have ratios δ_j, δ_k/δ_j, respectively.

Sequences with ratios resemble geometric progressions. If $\{b^i\}$ has ratio κ, then for an arbitrary $\epsilon>0$, with $\epsilon<\kappa$, there are numbers $T_1>0$, $T_2>0$ such that

$$T_1(\kappa-\epsilon)^i \leq |b^i| \leq T_2(\kappa+\epsilon)^i.$$

Thus, if $\kappa<1$, then b^i tends to zero more rapidly, eventually, than a geometric progression with ratio $\kappa+\epsilon$, and more slowly than the progression with ratio $\kappa-\epsilon$. Accordingly the ratio of a sequence is a measure of the speed of convergence to zero. We shall express our results on rate of convergence in these terms. Notice that when $\kappa<1$ the sequence $\{b^i\}$ is monotonic; decreasing if the sequence is positive, increasing if negative.

Suppose $\{b^i\}$, $\{c^i\}$ have ratios κ_1, κ_2 respectively. Then clearly $\{b^ic^i\}$ has ratio $\kappa_1\kappa_2$ and $\{b^i/c^i\}$ has ratio κ_1/κ_2. A convergent sequence $\{d^i\}$ with a nonzero limit has ratio 1. Consequently each of the sequences $\{b^id^i\}$, $\{b^i/d^i\}$ has ratio κ_1. The following lemma will be used frequently.

Lemma 6.2. Let $\{b^i\}$ have ratio κ. Suppose $\{d^i\}$ is a sequence such that d^i/b^i has a nonzero limit. Then $\{d^i\}$ has ratio κ.

For, d^i/b^i has ratio 1, and the product $d^i=(d^i/b^i)b^i$ has ratio κ.

If $\kappa\neq 1$, then the sequence of differences b^i-b^{i+1} also has ratio κ. For this we apply the last lemma with $d^i=b^i-b^{i+1}$.

Lemma 6.3. Suppose $\{b^i\}$ has ratio $\delta<1$. Let

$$B^i = \sum_{j=i}^{\infty} b^j.$$

Then

$$\lim_{i\to\infty} \frac{B^i}{b^i} = \frac{1}{1-\delta}.$$

In particular, $\{B^i\}$ has ratio δ.

Clearly we may assume that $\{b^i\}$ is positive. Select an arbitrary ϵ with $0<\epsilon<\delta$, $\delta+\epsilon<1$. Then for i sufficiently large, $(\delta-\epsilon)b^i \leq b^{i+1} \leq (\delta+\epsilon)b^i$, $(\delta-\epsilon)^2b^i \leq b^{i+2} \leq (\delta+\epsilon)^2b^i$, and in general

$$(\delta-\epsilon)^jb^i \leq b^{i+j} \leq (\delta+\epsilon)^jb^i, \qquad j=0,1,2,\ldots.$$

Hence, summing,

$$\frac{1}{1-\delta+\epsilon} \leq \frac{B^i}{b^i} \leq \frac{1}{1-\delta-\epsilon}$$

$$\frac{1}{1-\delta+\epsilon} \leq \underline{\lim}\, \frac{B^i}{b^i} \leqq \overline{\lim}\, \frac{B^i}{b^i} \leq \frac{1}{1-\delta-\epsilon}.$$

Since ϵ is arbitrary, the desired conclusion follows.

Lemma 6.4. Suppose $\{b^i\}$ is a positive sequence with ratio $\delta<1$. Let

$$Q^i = \prod_{j=i}^{\infty} (1+b^j).$$

Then

$$\lim_{i\to\infty} \frac{Q^i-1}{b^i} = \frac{1}{1-\delta}.$$

The existence of the infinite product Q^i follows from the fact that the series $\sum_{j=i}^{\infty} b^j$ of positive terms converges. Now

$$\prod_{i}^{i+l}(1+b^j) = 1+(b^i+b^{i+1}+\ldots+b^{i+l}) + (b^ib^{i+1}+\ldots)+\ldots+(b^ib^{i+1}\ldots b^{i+l}).$$

When i is sufficiently large for b^j, $j\geq i$, to be positive,

$$\sum_{i}^{i+l} b^j \leq \prod_{i}^{i+l}(1+b^j)-1 \leq B^i+(B^i)^2+ \ldots .$$

Allowing $l \to \infty$, we obtain

$$B^i \leq Q^i-1 \leq \frac{B^i}{1-B^i}.$$

Dividing by b^i, letting $i \to \infty$, and using lemma 6.3 we obtain the desired result.

We are now ready to state the main results of the section.

Theorem 6.1. Let eq (23) hold. Then $\mu^i-\lambda_1$ is a positive sequence with ratio δ_2^2.

This result is an immediate consequence of lemmas 5.3, 6.1, and 6.2.

From this theorem it follows that $\mu^i-\mu^{i+1}$ is also a sequence with ratio δ_2^2. Likewise, t^{i2} is a sequence with the same ratio, as we see from lemma 5.3.

Lemma 6.5. The positive sequence $L-a_1^i$ has ratio δ_2^2.

The number L is that of theorem 5.1. We have $a_1^i \to L$. From eq (16),

$$a_1^i=a_1^0 \prod_{k=0}^{i-1}\{1+\alpha(\mu^k-\lambda_1)\}, \quad L=a_1^0 \prod_{k=0}^{\infty}\{1+\alpha(\mu^k-\lambda_1)\}.$$

Thus

$$L-a_1^i=a_1^i\left[\prod_{k=i}^{\infty}\{1+\alpha(\mu^k-\lambda_1)\}-1\right].$$

Since $\mu^k-\lambda_1$ has ratio δ_2^2, the conclusion follows from lemma 6.4.

Theorem 6.2. Let eq (23) hold. Then the sequence $|x^i-Ly_1|$ has ratio δ_2.

From (16)

$$x-Ly_1=(a_1^i-L)y_1+a_2^iy_2+ \ldots +a_r^iy_r. \qquad (26)$$

Thus

$$|x-Ly_1|^2=(a_1^i-L)^2+a_2^{i\,2}+ \ldots +a_r^{i\,2}.$$

Hence the left side is a sum of terms having ratios $\delta_2^2, \delta_2^2, \delta_3^2, \ldots, \delta_r^2$. Since each of these numbers, other than the second, is strictly smaller than the second, the theorem is established.

Concerning the convergence of $|x^i|$ to L, we remark that $L-|x^i|$ is a sequence of ratio δ_2^2. The proof is like that of lemma 6.5, with a_1^i replaced by $|x^i|^2$.

Information on the convergence of each component of x^i can be derived from eq (26). We write $x^i=(b_1^i, b_2^i, \cdots, b_n^i)$, $y_1=(c_1, c_2, \cdots, c_n)$. Now for fixed j, $b_j^i-Lc_j$ is a linear combination of $a_1^i-L, a_2^i, \cdots, a_r^i$, these sequences having respective ratios $\delta_2^2, \delta_2, \delta_3, \cdots, \delta_r$. If the j^{th} component of y_2, the coefficient of a_2^i, is not 0, then $b_j^i-Lc_j$ has ratio δ_2; accordingly the difference sequence $b_j^i-b_j^{i+1}$, which is numerically available, has ratio δ_2. If the j^{th} component of y_2 is 0, then, in general, $b_j^i-Lc_j$ will have a smaller ratio (assuming the sequence is not identically 0).

For results on the convergence of sequences associated with the second characteristic vector y_2, we appeal to the forthcoming theorems of section 11. These results may be specialized to the sequence $\{\xi^i\}$ by recognizing that in the notation of that section,

$$x^i=z_1^i, \quad \xi^i=z_2^i, \quad \mu_1^i=\mu(x^i), \quad \mu_2^i=\mu(\xi^i).$$

Also, we agree to interpet δ_{r+1} as 0.

Concerning the rate of convergence of $|\xi^i|$ to 0 we have, from lemma 5.3, that $\{|\xi^i|\}$ is a sequence with ratio δ_2.

Theorem 6.3. Let eq (23) hold and set $\eta^i=\xi^i/|\xi^i|$.

(1) If $\delta_2^2>\delta_3$, then $\{\lambda_2-\mu(\xi^i)\}$ is a positive, monotonic sequence with ratio δ_2^2, and $\{|\eta^i-y_2|\}$ is a sequence with ratio δ_2.

(2) If $\delta_2^2<\delta_3$, then $\{\mu(\xi^i)-\lambda_2\}$ is a positive, monotonic sequence with ratio $(\delta_3/\delta_2)^2$, and $\{|\eta^i-y_2|\}$ is a sequence with ratio δ_3/δ_2.

(3) If $\delta_2^2=\delta_3$, then $\mu(\xi^i)-\lambda_2=\mathbf{O}(\delta_2^2)$ and $|\eta^i-y_2|=\mathbf{O}(\delta_2)$.

These statements are interpretations of theorems 11.1, 11.2, and their corollaries for $j=2$. From theorem 11.4 we derive the inequality

$$\lambda_2-\mu(\xi^i) \leq \mu(x^i)-\lambda_1, \quad i=0, 1, 2, \cdots,$$

the equality holding just in case $r=2$. From this and the above theorem we obtain the following interesting corollary.

Corollary. If $\delta_2^2>\delta_3$, then for sufficiently large i,

$$|\lambda_2-\mu(\xi^i)| \leq |\lambda_1-\mu(x^i)|.$$

More generally, when $\lambda_2-\mu(\xi^i)$ is nonnegative, and i is large, then $\mu(\xi^i)$ is closer to λ_2 than $\mu(x^i)$ is to λ_1. On the other hand, for case (2) of theorem 6.3, the sequence $\{\mu(\xi^i)-\lambda_2\}$ has ratio $(\delta_3/\delta_2)^2>\delta_2^2$. But δ_2^2 is the ratio of $\{\mu^i-\lambda_2\}$. Hence in this instance μ^i approaches λ_1 more rapidly than $\mu(\xi^i)$ approaches λ_2.

Notice that for $r=2$, the inequality $\delta_2^2>\delta_3$ is automatic. If $r>2$, then, from theorem 11.3, the inequality $\delta_2^2>\delta_3$ holds whenever

$$\lambda_3-\lambda_2>\lambda_2-\lambda_1.$$

VII. The Method of Optimum a

In sections 10 and 11 the method of fixed α will be extended to obtain sequences that converge to all of the characteristic vectors $y_1, y_2, \cdots, y_r$. Before proceeding to this generalization we wish to develop first the method of optimum α, so that a comparison of the two methods (and others) may be made at the most advantageous point. The optimum pro-

cedure does not appear to generalize to a larger number of characteristic vectors as simply as the method of fixed α.

According to eq (10) the iteration scheme is now given by eq (12) with

$$\gamma^i=\frac{2}{s^i+\sqrt{s^{i2}+4t^{i2}}}>0. \tag{27}$$

The numbers t^i are bounded (see remarks preceding theorem 5.1); so are the numbers s^i. Thus $f(\gamma^i)=\gamma^i$ is bounded from below by a positive constant. Accordingly, theorem 4.1 yields the following lemma.

Lemma 7.1. Let γ^i be given by eq (31). Then the conclusion of theorem 4.1 holds.

In order to apply theorem 4.2 it is necessary to establish the boundedness of the γ^i. This is done in the next lemma.

Lemma 7.2. Let γ^i be given by eq (31). Then

$$\overline{\lim_{i\to\infty}}\,\gamma^i\leqq\frac{1}{\lambda_2-\lambda_1}.$$

Since $\mu^i\to\lambda_1$ it is sufficient, by eq (11), to show that

$$\underline{\lim_{i\to\infty}}\,\mu(\xi^i)\geq\lambda_2.$$

Let T denote the inferior limit on the left. Let

$$\eta^i=\frac{\xi^i}{|\xi^i|}.$$

Then $T=\lim\,\mu(\eta^i)$. Select a subsequence $\{\eta^k\}$ such that $T=\overline{\lim}\,\mu(\eta^k)$. Select a further subsequence $\{\eta^l\}$ that converges to, say, η. From $(\eta^i, x^i)=0$ and lemma 7.1 we obtain

$$0=\lim_{l\to\infty}\left(\eta^l,\frac{x^l}{|x^l|}\right)=(\eta, y_1).$$

Since η belongs to $\mathscr{A}$ (space spanned by $y_1,\cdots,y_r$) it follows that $\mu(\eta)\geq\lambda_2$. Hence

$$T=\lim_{l\to\infty}\mu(\eta^l)=\mu(\eta)\geq\lambda_2,$$

as desired.

Theorem 4.2 now yields the principal result.

Theorem 7.1. Let γ^i be given by eq (27). Then

$$\lim_{i\to\infty}x^i=Ly_1,\quad\lim_{i\to\infty}\mu^i\to\lambda_1.$$

That the vectors $\xi^i/|\xi^i|$ do not converge in the method of optimum α is a consequence of the next theorem.

Theorem 7.2. Under the hypotheses of theorem 7.1,

$$(\xi^i,\xi^{i+1})=0.$$

Proof. We have,

$$\begin{aligned}\xi^{i+1}&=Ax^{i+1}-\mu^{i+1}x^{i+1}\\&=Ax^i-\gamma^iA\xi^i-\mu^{i+1}x^i+\mu^{i+1}\gamma^i\xi^i.\end{aligned}$$

Hence

$$\begin{aligned}(\xi^i,\xi^{i+1})&=(\xi^i,Ax^i)-\gamma^i(A\xi^i,\xi^i)+\mu^{i+1}\gamma^i(\xi^i,\xi^i)\\&=\{1-\gamma^i(\mu(\xi^i)-\mu^{i+1})\}(\xi^i,\xi^i)\\&=0\end{aligned}$$

by eq (11).

The last theorem indicates that the coefficients a_j^i of eq (16) converge in an irregular fashion as $i\to\infty$. Hence a determination of the rate of convergence such as that given in section 5 for the method of fixed α is not to be expected here.

We conclude the section with a modification of theorem 7.1 for convergence to the highest characteristic vector y_r.

Theorem 7.3. Let the sequence $\{x^i\}$ be defined by

$$x^{i+1}=x^i+\gamma^i\xi^i,$$

with

$$\gamma^i=\frac{2}{-s^i+\sqrt{s^{i2}+4t^{i2}}}.$$

Then

$$\lim_{i\to\infty}x^i\to Ly_r,\quad L>0.$$

This result is obtained by applying theorem 7.1 to the matrix $B=-A$.

VIII. Comparison of Methods

Each step in the preceding methods of gradients consists of forming the subspace spanned by x and ξ, i. e., by x and Ax, and choosing in that two-space[6] a next approximation $\overline{x}$. In the method of optimum α, the vector $\overline{x}$ is chosen so that $\mu(\overline{x})$ is a minimum in the subspace; in the method of fixed α the particular linear combination $x-\alpha\xi$ is chosen. Superficially, then, it would seem that the first method should give more rapid convergence to λ_1 and y_1 than the second. But the minimizing procedure can only be recognized as an advantage when the vector x^i is common to the two methods. This certainly holds for the initial vector x^0, but not necessarily beyond this stage; hence the relative merits of the two methods are not evident. It seems reasonable that for low order matrices the procedure of finding an optimum α might be advantageous, while for matrices of higher order the fixed α procedure might be superior. In practice a combination of the two methods would be in order.

The method of fixed α has the advantage of computational simplicity over the alternative method. On the other hand the former requires some advance

[6] We shall remark later on the possibility of dealing with subspaces of dimension higher than two. Cf. a forthcoming paper by W. Karush, An iterative method for finding characteristic vectors of a symmetric matrix. There the method of optimum α is generalized to subspace of arbitrary (fixed) dimension.

information on the characteristic roots in order to estimate an allowable α; computation by the method of optimum α demands no such knowledge.

The method of fixed α is a smooth method. The vectors $\xi^i/|\xi^i|$ converge; by a relatively simple extension of the procedure we can obtain convergence to all the characteristic vectors of $\mathscr{A}_r$ (see section 10). Furthermore the convergence is geometric in nature. The method of optimum α has none of these advantages. The successive gradients are orthogonal, and the coefficients a_2^i, a_3^i, . . . of eq (16) do not tend to 0 smoothly.

Sequences $\{b^i\}$ with ratio $\delta<1$ provide examples of "linear" convergence. If the ratio $b^{i+1}/(b^i)^2$ has a finite limit, then the sequence converges in a "quadratic" fashion. In the method of fixed α we have convergence of a linear type. It is possible to procure quadratic convergence by modifying the method, but the price paid would be the solution of a system of linear equations at each step of the iteration.

A combination of the two gradient methods can be used to advantage when the characteristic roots λ_1, λ_2 are close and relatively isolated from the other roots. Suppose we begin with the fixed α iteration and continue up to a certain stage. The last estimate x and its gradient ξ may not be good individual approximations to y_1 and y_2 but the pair may provide an excellent approximation to the plane of y_1 and y_2. One application of the optimum α procedure at this stage should then yield a good approximation $\bar{x}$ to y_1. The iteration may be started again with $\bar{x}$. This technique has worked well in several numerical examples.

Since $t=|\xi|/|x|\to 0$, it is clear that fewer and fewer significant figures will be retained in ξ as the iteration proceeds in a numerical calculation. Nevertheless, ξ and $\mu(\xi)$ supply useful approximations to y_2 and λ_2. For example, after y_1 and λ_1 have been obtained, the approximation ξ can be used as the initial vector x^0 in the calculation of y_2 and λ_2 by a new use of the method of gradients.

The above remarks apply equally well to the highest vector y_r. One of the advantages of the gradient methods is that they can be applied directly to either end of the scale of characteristic roots.

A single step (2) of the gradient procedure carries one from

$$x=a_1y_1+a_2y_2+\cdots+a_ry_r$$

to

$$\bar{x}=\{1+\alpha(\mu-\lambda_1)\}a_1y_1+\{1+\alpha(\mu-\lambda_2)\}a_2y_2+\cdots.$$

Thus if any characteristic number λ_j is known, then orthogonality of $\bar{x}$ to y_j can be achieved by choosing[7]

$$\alpha=\frac{1}{\lambda_j-\mu}.$$

[7] The same result can be obtained by taking $\bar{x}=Ax-\lambda_j x$; the advantage of the above procedure is that it fits into the iteration scheme (12).

This observation may be used to maintain (approximate) orthogonality which may be theoretically assured but may be gradually lost in computation due to round-off errors. For example, in beginning with ξ as an initial vector for obtaining y_2 (see second paragraph above), one must maintain orthogonality to y_1 during the iteration process. To do this, the procedure (12) may be interspersed with steps in which $\gamma^i=(\lambda_1-\mu^i)^{-1}$. Similarly, one may maintain orthogonality to any number of characteristic vectors, or induce orthogonality to such vectors, provided only that one knows the corresponding characteristic roots (relatively accurately). The desired ends are achieved within the framework of the scheme (12) by particular choices of γ^i. Notice that if $r-1$ of the r characteristic roots λ_j are known then the characteristic vector belonging to the remaining root can be obtained in $r-1$ steps.

It was remarked after theorem 6.1 that $\Delta\mu^i=\mu^{i+1}-\mu^i$ and t^i are sequences with ratio δ_2^2. Hence observation of the ratios $\Delta\mu^{i+1}/\Delta\mu^i$ or t^{i+1}/t^i can lead to an estimate of δ_2^2. Since $\delta_2=1-\alpha(\lambda_2-\lambda_1)$, and α is known, another estimate of λ_2 may be made, alternative to the estimate $\mu(\xi)$. Furthermore, if a sequence c^i has a limit C, and the difference c^i-C is a sequence with a ratio $\delta<1$, then a transformation for speeding the convergence of c^i is

$$\bar{c}^i=c^i+\frac{\Delta c^i}{1-\delta^i}, \qquad \left(\delta^i=\frac{\Delta c^{i+1}}{\Delta c^i}\right)$$

$$=\frac{c^{i+1}-\delta^i c^i}{1-\delta^i}$$

$$\approx\frac{c^{i+1}-\delta c^i}{1-\delta}.$$

With an estimate of δ_2, this formula may be used to improve μ^i and the components of x^i (see remarks following theorem 6.2).

A certain degree of flexibility may be introduced in the method of fixed α by replacing the constant α by a variable α^i,

$$\alpha^i=\frac{\beta^i}{M}, \tag{28}$$

which satisfies

$$0<\beta_1\leq\beta^i\leq\beta_2<2, \qquad (\beta_1,\ \beta_2 \text{ constants}). \tag{29}$$

The conclusion of theorem 5.1 remains valid, since this choice of α^i does not violate the assumptions of theorem 4.2.

The convergence of the gradients as in theorem 5.2 may also be justified provided one imposes the additional restrictions

$$\beta_2<1 \text{ and } \frac{\beta_2}{\beta_1}<\frac{\lambda_3-\lambda_1}{\lambda_2-\lambda_1}. \tag{30}$$

Under these conditions statements on rate of con-

vergence, appropriately formulated, can be established.

For comparison purposes we state here, without proof, some properties of the "power" method. The iteration formula is

$$x^{i+1}=Ax^i.$$

Suppose now that there is one characteristic root of maximum absolute value; denote it by λ_k. Then

$$\lim_{i\to\infty}\frac{x^{2i}}{|x^{2i}|}\to y_k, \qquad \lim_{i\to\infty}\mu(x^i)\to\lambda_k.$$

(In the first limit the sequence with odd indices tends to $-y_k$ if $\lambda_k<0$, and to y_k if $\lambda_k>0$.) It was pointed out in the introduction that the power method is a gradient method with

$$\gamma^i=-\frac{1}{\mu^i}.$$

In this form the iteration formula is

$$z^{i+1}=\frac{1}{\mu^i}Az^i, \quad \mu^i=\mu(z^i),$$

$$\text{i. e., } z^{i+1}=(\mu^0\mu^1\cdots\mu^i)^{-1}x^{i+1},$$

and the convergence result is

$$\lim_{i\to\infty}z^i=K\,y_k, \qquad K=\text{const.}$$

(provided μ^i is never 0). Thus μ^i, as well as $|x^i|$, may be used as a normalization factor.

The power method leads directly to an estimate of λ_k, and hence to an upper bound $2|\lambda_k|$ of M. Thus a rough application of this method may be used as a preliminary step in the method of fixed α in order to determine an allowable value of the constant α.

Suppose now that λ_l is a characteristic root of next highest absolute value. Concerning rate of convergence we remark that $\mu^i-\lambda_k$ is a sequence of ratio $(\lambda_l/\lambda_k)^2$. If $-\lambda_l$ is not a characteristic root, then $\xi^{2i}=\xi(x^{2i})$ tends in direction to

$$\frac{(\lambda_l-\lambda_k)}{|\lambda_l-\lambda_k|}\,y_l;$$

the sequence with odd indices tends to the negative of this vector if $\lambda_l<0$ otherwise to the same vector.

We conclude this section with a remark on the method of relaxations.[8] Any method which begins with a vector x and applies a sequence of transformations to it so that the resulting sequence of vectors x^i converges to a characteristic vector, must induce the gradients $\xi(x^i)$ to tend to the null vector. Thus any artful modification of a vector x to a new vector $\bar{x}$ which brings the gradient closer to the null vector would be a plausible procedure in a numerical calculation. The full skill and intuition of the computer may come into play in varying the vector at any stage to produce a better approximation. One systematic procedure is to modify a single component of x in such a way as to make a single component of the gradient vanish. Other devices may be used; the best trick at any point depends upon the information available at that stage. This is the flexible approach of the method of relaxations. Clearly it is not easily adaptable to automatic computation.

[8] Cooper, J. L. B., The Solution of Natural Frequency Equations by Relaxation Methods, Qtrly. Appl. Math., vol. 6, p. 179 (1948).

IX. Invariant Subspace $\mathscr{A}_r$

Before turning to the generalization of the method of fixed α we shall develop here some properties of the invariant subspace $\mathscr{A}_r$. We will encounter certain polynomials $p_j(\lambda)$ which have been considered by Lanczos,[9] and some of the results of this section will overlap his work.

In place of the notation x^0 we shall use z_1 to denote a given vector with expansion

$$z_1=a_1y_1+a_2y_2+\cdots+a_ry_r \tag{31}$$

as in eq (4). The space $\mathscr{A}_r=(y_1\cdots,y_r)$ may be characterized as in the following lemma.

Lemma 9.1. The vectors $z_1, Az_1, \cdots, A^{r-1}z_1$ span the space $\mathscr{A}_r$.

Let $\mathscr{B}$ be the space spanned by the vectors A^kz_1, $k=0, 1, 2, \cdots$. Then this space is the smallest invariant subspace which contains the vector z_1. Since z_1 is in $\mathscr{A}_r$, and $\mathscr{A}_r$ is invariant, it follows that the relation $\mathscr{B}<\mathscr{A}_r$ holds.

Suppose that $\mathscr{B}$ were a proper subset of the invariant space $\mathscr{A}_r$. Then $\mathscr{B}$ would be spanned by a proper subset of the characteristic vectors $y_1, y_2, \cdots, y_r$ of $\mathscr{A}_r$. Consequently the vector z_1, being in $\mathscr{B}$, would be orthogonal to at least one of these characteristic vectors, say y_k. Thus, from eq (31), $a_k=(z_1,y_k)=0$. This contradicts the assumption $a_k>0$ (see eq 4).

We now have $\mathscr{B}=\mathscr{A}_r$. Consequently, dim $\mathscr{B}$=dim $\mathscr{A}_r=r$. The lemma now follows from the fact that dim $\mathscr{B}$ is the least integer h such that the vectors z_1, $Az_1, \ldots, A^{h-1}z_1$ span $\mathscr{B}$.

We now define the vectors z_j recursively as follows. Let

$$\left.\begin{aligned} t_j&=\frac{|z_j|}{|z_{j-1}|}\\ \mu_j&=\mu(z_j).\end{aligned}\right\} \tag{32}$$

Then

$$\left.\begin{aligned} z_2&=Az_1-\mu_1z_1\\ z_3&=Az_2-\mu_2z_2-t_2^2z_1\\ &\cdots\cdots\cdots\cdots\\ z_{j+1}&=Az_j-\mu_jz_j-t_j^2z_{j-1}, \quad j=2, 3, \ldots, r.\end{aligned}\right\} \tag{33}$$

[9] "An Iteration Method for the Solution of the Eigenvalue Problem of Linear Differential and Integral Operators," J. Research NBS 45, 255 (1950) RP 2133

The vectors $z_1, z_2, \ldots, z_r, z_{r+1}$ are well-defined and

$$z_j \neq 0, \qquad j=1,2, \ldots, r.$$

For suppose, $z_1, \ldots, z_k$ are defined and none is null, $k<r$. Then by eq (33) z_{k+1} is defined and is a linear combination of $A^k z_1, A^{k-1} z_1, \ldots, A z_1, z_1$ with the coefficient of $A^k z_1$ equal to one. Further, if $k<r-1$, then $z_{k+1} \neq 0$ by lemma 9.1. This establishes the assertion.

By induction it is easy to prove that

$$(z_j, z_j)=(z_j, A z_{j-1}), \qquad j=2, \ldots, r+1,$$

$$(z_j, z_k)=0, \qquad j \neq k, \quad j,k=1,2, \ldots, r+1. \tag{34}$$

Lemma 9.2. Let $\mathscr{A}_j$ be the space spanned by $z_1, A z_1, \ldots, A^{j-1} z_1$, i. e.,

$$\mathscr{A}_j=(z_1, A z_1, \ldots, A^{j-1} z_1), \qquad j=1, \ldots, r.$$

Then $(z_1, \ldots, z_j)$ is an orthogonal basis for $\mathscr{A}_j$. Also,

$$z_{r+1}=0.$$

By lemma (9.1), $\mathscr{A}_j$ has dimension j. From the above, the nonnull vectors $(z_1, \ldots, z_j)$ are orthogonal, and lie in $\mathscr{A}_j$; hence they span $\mathscr{A}_j$.

The final conclusion of the lemma follows from the fact that z_{r+1} is orthogonal to $z_1, \ldots, z_r$ and hence to $\mathscr{A}_r$. But from the invariance of $\mathscr{A}_r$, z_{r+1} belongs to $\mathscr{A}_r$. Thus $z_{r+1}=0$.

Given an arbitrary vector z_1 the integer r of eq (31) can be determined from the above lemma; r is that integer such that z_{r+1} is the first z_j that vanishes. Notice that except for $j=r$, $\mathscr{A}_j$ is *not* the space spanned by $y_1, \ldots, y_j$.

We consider now the expansions of the z_j in terms of the y_j. From eq (31) and (33),

$$z_2=(\lambda_1-\mu_1)\, a_1 y_1+(\lambda_2-\mu_1)\, a_2 y_2+ \cdots +(\lambda_r-\mu_1)\, a_r y_r.$$

Thus, if we put $p_1(\lambda)=\lambda-\mu_1$ we have

$$z_2=p_1(\lambda_1)\, a_1 y_1+p_1(\lambda_2)\, a_2 y_2+ \cdots + p_1(\lambda_r)\, a_r y_r.$$

Again from eq (33),

$$z_3=\{p_1(\lambda_1)(\lambda_1-\mu_2)-t_2^2\} a_1 y_1+ \cdots + \{p_1(\lambda_r)(\lambda_r-\mu_2)-t_2^2\} a_r y_r.$$

Hence, if we put $p_2(\lambda)=p_1(\lambda)(\lambda-\mu_2)-t_2^2$ we have

$$z_3=p_2(\lambda_1)\, a_1 y_1+p_2(\lambda_2)\, a_2 y_2+ \cdots + p_2(\lambda_r)\, a_r y_r.$$

In general, if we define the polynomials $p_j(\lambda)$ by

$$p_0(\lambda)=1, p_1(\lambda)=\lambda-\mu_1, p_j(\lambda)=p_{j-1}(\lambda)(\lambda-\mu_j) - t_j^2 p_{j-2}(\lambda), j=2, \cdots, r \tag{35}$$

we have the expansions

$$z_{j+1}=p_j(\lambda_1)\, a_1 y_1+p_j(\lambda_2)\, a_2 y_2+ \cdots + p_j(\lambda_r)\, a_r y_r, \quad j=0, 1, \cdots, r. \tag{36}$$

From eq (36) with $j=r+1$, and lemma 9.2 we see that the roots of

$$p_r(\lambda)=0$$

are the characteristic numbers $\lambda_1, \lambda_2, \cdots, \lambda_r$. This is a special case of the forthcoming theorem 9.2.

In lemma 9.2 we defined the subspaces $\mathscr{A}_j$ of $\mathscr{A}_r$. In order to develop some properties of these subspaces we state without proof several well known properties of symmetric linear operators.

Let $\mathscr{B}$ denote a linear space of real vectors. Let B denote a symmetric linear operator on $\mathscr{B}$. Consider an l-dimensional subspace $\mathscr{B}'$ of $\mathscr{B}$; let π be the projection operator into $\mathscr{B}'$. The operator πB with domain restricted to $\mathscr{B}'$ is a symmetric operator on $\mathscr{B}'$; by the characteristic roots and vectors of B relative to $\mathscr{B}'$ we mean the corresponding quantities of this operator. An alternative characterization is given by the property that v is a characteristic vector (relative to $\mathscr{B}'$) belonging to the characteristic root κ if and only if v belongs to $\mathscr{B}'$, $v \neq 0$, and

$$(x, Bv)=\kappa(x, v), \qquad x \text{ in } \mathscr{B}'.$$

Also, the smallest and greatest roots relative to $\mathscr{B}'$ are respectively the minimum and maximum of $\mu(x)$ on $\mathscr{B}'$. If $\mathscr{B}'$ is an invariant subspace of $\mathscr{B}$ then $\pi B=B$ on $\mathscr{B}'$ and the characteristic roots and vectors relative to $\mathscr{B}'$ are such in the whole space $\mathscr{B}$; this is the case with $\mathscr{A}_r$ as a subspace of $\mathscr{A}$.

If $u_1, u_2, \cdots, u_l$ is an orthonormal basis for $\mathscr{B}'$ then the operator πB relative to this basis has the $l \times l$ matrix representation

$$B'=(A u_j, u_k), \qquad j,k=1,2, \cdots, l.$$

The characteristic roots relative to $\mathscr{B}'$ are the roots of

$$|\lambda I-B'|=0.$$

We quote a well-known result. (The first conclusion of the theorem below holds without the restriction that the κ_j be distinct.)

Theorem 9.1. Let B be a symmetric linear operator on an m-dimensional real linear space $\mathscr{B}$; let $\mathscr{B}'$ be an l-dimensional subspace of $\mathscr{B}$. Suppose that B has distinct characteristic values $\kappa_1<\kappa_2<\cdots<\kappa_m$ with corresponding characteristic vectors $v_1, v_2, \cdots, v_m$. Let the characteristic roots and vectors of B relative to $\mathscr{B}'$ be respectively $\kappa_1' \leq \kappa_2' \leq \cdots \leq \kappa_l'$ and $v', v_2', \cdots, v_l'$. Then

$$\kappa_1 \leq \kappa_1', \quad \kappa_2 \leq \kappa_2', \cdots, \quad \kappa_l \leq \kappa_l'.$$

The equality holds for all indices if and only if

$$v_1=v_1', \quad v_2=v_2', \ \ldots\ , v_i=v_i';$$

in this case $\mathscr{B}'$ is an invariant subspace.[10]

We return to the subspaces $\mathscr{A}_j$. The characteristic roots and vectors of A relative to $\mathscr{A}_j$ are those of the symmetric operator $B_j=\pi_j A$ on $\mathscr{A}_j$, where π_j is the projection on $\mathscr{A}_j$. Since $z_1, \ldots, z_j$ is an orthogonal basis of $\mathscr{A}_j$ the operator B_j has the matrix representation

$$\left(\frac{(Az_l, z_m)}{|z_l|\,|z_m|}\right), \qquad l, m=1, 2, \ldots, j.$$

We calculate this matrix.

Lemma 9.3. For a given j,

$$\begin{aligned}(Az_j, z_k)&=|z_j|^2 && \text{when } k=j-1\\ &=\mu_j|z_j|^2 && \text{when } k=j\\ &=|z_{j+1}|^2 && \text{when } k=j+1\\ &=0 && \text{otherwise.}\end{aligned}$$

We remark that for $j=1$, the first equation is to be omitted. The first and third equations follow from the first equation of (34). The second follows from the definition of μ_j. The last is a consequence of the orthogonality relations of (34) and the fact that Az_j is a linear combination of z_{j+1}, z_j and z_{j-1} by (33).

It follows from this lemma that B_j has the matrix representation

$$B_j=\begin{pmatrix} \mu_1 & t_2 & 0 & & & \\ t_2 & \mu_2 & t_3 & \cdot & & \\ 0 & t_3 & \mu_3 & \cdot & \cdot & \\ & \cdot & \cdot & \cdot & \cdot & \cdot \\ & & \cdot & \cdot & \cdot & 0 \\ & 0 & & \cdot & \cdot & t_j \\ & & & 0 & t_j & \mu_j \end{pmatrix}$$

Lemma 9.4. Let $q_0(\lambda)=1$ and $q_j(\lambda)$ be the characteristic polynomial of B_j,

$$q_j(\lambda)=|\lambda I-B_j|, \qquad j=1, 2, \cdots, r.$$

Then

$$p_j(\lambda)=q_j(\lambda), \qquad j=0, 1, \cdots, r.$$

Using the matrix representation above it is easy to see by direct calculation that

$$q_1(\lambda)=\lambda-\mu_1, \quad q_j(\lambda)=(\lambda-\mu_j)\,q_{j-1}(\lambda)-t_j^2\,q_{j-2}(\lambda).$$

The lemma now follows from the fact that these recursion relations are identical with eq (35).

Theorem 9.2. The roots $\nu_1\leq\nu_2\leq \ldots \leq\nu_j$ of

$$p_j(\lambda)=0, \ j=1, 2, \ldots, r,$$

are the characteristic roots of A relative to the subspace $\mathscr{A}_j$. Furthermore,

$$\lambda_1\leq\nu_1, \ \lambda_2\leq\nu_2, \ \ldots, \ \lambda_j\leq\nu_j,$$

the equality holding for all indices if and only if $j=r$.

This theorem follows from theorem 9.1 with $\mathscr{B}=\mathscr{A}_j$, $B=A$. The last statement follows from the fact that $\mathscr{A}_j$, $j<r$, is not invariant.

X. Extension of Method of Fixed α

We return now to the sequence $\{x^i\}$ of eq (12) with $\gamma^i=\alpha=$const. As in theorem 5.2 we impose the condition (23)[11] on α. We rename x^i as z_1^i and apply the results of the previous section to each z_1^i. The expansion (31) becomes (16),

$$z_1^i=a_1^i y_1+a_2^i y_2+ \ \ldots \ +a_r^i y_r.$$

By lemma 5.1 each a_j^i is positive; accordingly, the space $\mathscr{A}_r$ associated with z_1^i is independent of i and coincides with our original space $\mathscr{A}$. We recognize z_2^i as our originial ξ^i.

The polynomials of eq (35) now depend upon the index i; we denote them by $p_j^i(\lambda)$. For each i they are defined in terms of

$$\mu_j^i=\mu(z_j^i), \qquad j=1, 2, \ldots, r,$$

$$t_j^i=\frac{|z_j^i|}{|z_{j-1}^i|}, \qquad j=2, \ldots, r.$$

It is convenient to introduce

$$\begin{aligned}&\overline{p}_0(\lambda)=1,\\ &\overline{p}_j(\lambda)=(\lambda-\lambda_j)(\lambda-\lambda_{j-1})\cdots(\lambda-\lambda_1),\\ &\qquad j=1, 2, \cdots, r.\end{aligned}$$

The vectors $(z_1^i, z_2^i, \cdots, z_r^i)$ comprise a sequence of orthogonal bases for the space $\mathscr{A}_r$. We show that, when normalized, this sequence converges to the fixed basis $y_1, \cdots, y_r$.

Theorem 10.1. Let the constant α satisfy (23) and let the initial vector $z_1^0=x^0$ be given by (4). Determine the infinite sequences $\{z_1^i\}, \{z_2^i\}, \cdots, \{z_r^i\}$ by (33) and

$$z_1^{i+1}=z_1^i-\alpha z_2^i.$$

[10] This result follows from the minimax principle for characteristic values. Cf. Courant and Hilbert "Methoden der Mathematischen Physik," I, 2d ed., Berlin (1931), pp. 27–29.

[11] In this section and the next we assume without further comment that this condition holds.

Then

$$\lim_{i\to\infty}\frac{z_j^i}{z_1^i}=y_j,\quad \lim_{i\to\infty}\mu(z_j^i)=\lambda_j,\quad j=1,2,\cdots,r.$$

The proof will be made by induction. For an integer k, $2\leq k\leq r$, consider the statement

$$\lim_{i\to\infty}\frac{z_j^i}{|z_1^i|}=y_j,\quad \lim_{i\to\infty}\frac{|z_j^i|}{a_j^i}=\overline{p}_{j-1}(\lambda_j),\quad j=1,2,\cdots,k. \tag{37}$$

By theorem 5.1, (5.2), lemma 5.3 and theorem 5.2 we see that the statement is true for $k=2$. Assuming the statement true for $k<r$, we shall show that it holds for $k+1$. This will provide a proof of the theorem.

From the second eq (37), the definition of t_j^i, and lemma 5.2 it follows that

$$t_j^i\to 0,\qquad j=2,\cdots,k.$$

Also by the first eq (37)

$$\mu_j^i\to\lambda_j,\qquad j=1,2,\cdots,k.$$

Consequently, from eq (35), $p_1^i(\lambda)\to\lambda-\lambda_1=\overline{p}_1(\lambda)$, $p_2^i\to(\lambda-\lambda_2)\overline{p}_1(\lambda)=\overline{p}_2(\lambda)$, and generally,

$$p_j^i(\lambda)\to\overline{p}_j(\lambda),\qquad j=1,2,\cdots,k.$$

The vector under examination is

$$z_{k+1}^i=p_k^i(\lambda_1)a_1^iy_1+p_k^i(\lambda_2)a_2^iy_2+\cdots+p_k^i(\lambda_r)a_r^iy_r. \tag{38}$$

From the orthogonality relation $(z_j^i,z_{k+1}^i)=0$ and eq (36),

$$\sum_{j=1}^{r}p_0^i(\lambda_j)p_k^i(\lambda_j)a_j^{i\,2}=0$$

$$\cdots\cdots\cdots\cdots$$

$$\sum_{j=1}^{r}p_{k-1}^i(\lambda_j)\,p_k^i(\lambda_j)\,a_j^{i\,2}=0.$$

Dividing by $a_{k+1}^{i\,2}$ and writing

$$\alpha_j^i=\frac{p_k^i(\lambda_j)\,a_j^{i\,2}}{a_{k+1}^{i\,2}},\quad j=1,2,\cdots,k,$$

these equations take the form

$$p_0^i(\lambda_1)\,\alpha_1^i+p_0^i(\lambda_2)\,\alpha_2^i+\cdots+p_0^i(\lambda_k)\,\alpha_k^i=b_0^i$$

$$\cdots\quad\cdots\quad\cdots\quad\cdots\quad\cdots\quad\cdots\quad\cdots$$

$$p_{k-1}^i(\lambda_1)\,\alpha_1^i+p_{k-1}^i(\lambda_2)\,\alpha_2^i+\cdots+p_{k-1}^i(\lambda_k)\,\alpha_k^i=b_{k-1}^i.$$

By lemma 5.2,

$$b_j^i\to\overline{p}_j(\lambda_{k+1})\,\overline{p}_k(\lambda_{k+1}).$$

As $i\to\infty$ the matrix of coefficients $(p_j^i(\lambda_l))$ tends to $(\overline{p}_j(\lambda_l))$. The latter matrix has only zeros below the main diagonal; its determinant is the product of the diagonal elements, namely,

$$\overline{p}_0(\lambda_1)\overline{p}_1(\lambda_2)\cdots\overline{p}_{k-1}(\lambda_k).$$

From the definition of the polynomials $\overline{p}_j$ and the fact that the λ_j are distinct, it follows that this quantity does not vanish. Consequently, for i sufficiently large, the above linear equations in α_j^i may be solved for these unknowns. Furthermore, as i tends to infinity, the solutions have finite limits, say L_j. Hence

$$\alpha_j^i\to L_j,\quad j=1,2,\ldots,k. \tag{39}$$

It is simple to compute L_k; allowing $i\to\infty$, we obtain, from the last equation,

$$L_k=-\frac{\overline{p}_k(\lambda_{k+1})\overline{p}_{k-1}(\lambda_{k+1})}{\overline{p}_{k-1}(\lambda_k)}.$$

For our purposes it is not essential to know more of the limiting values L_j than the fact that they exist. As a matter of interest they are evaluated at the end of the section.

Now divide both sides of eq (38) by a_{k+1}^i. From lemma 5.2 and eq (39),

$$\frac{z_{k+1}^i}{a_{k+1}^i}\to\overline{p}_k(\lambda_{k+1})\,y_{k+1}.$$

From this we immediately obtain the eq (37) for $j=k+1$. This completes the proof.

For future use we record some of the results of the above proof in the following corollary.

Corollary. For $j=0,1,\cdots,r$,

$$p_j^i(\lambda)\to\overline{p}_j(\lambda),\quad \frac{|z_j^i|}{a_j^i}\to\overline{p}_{j-1}(\lambda_j).$$

Also, for $j=1,2,\cdots,l-1$, $l=2,3,\cdots,r$,

$$\frac{p_{l-1}^i(\lambda_j)\,a_j^{i\,2}}{a_l^{i\,2}}$$

has a finite limit; this limit has the value

$$-\frac{\overline{p}_{l-1}(\lambda_l)\,\overline{p}_{l-2}(\lambda_l)}{\overline{p}_{l-2}(\lambda_{l-1})}$$

when

$$j=l-1.$$

We return to the evaluation of L_j of eq (39). Let

$$\beta_j=\frac{L_j}{\overline{p}_k(\lambda_{k+1})}. \tag{40}$$

Then the β_j satisfy the limiting equations

$$\beta_1+ \quad \beta_2+\ldots+ \quad \beta_k=-1$$

$$\overline{p}_1(\lambda_1)\beta_1+\overline{p}_1(\lambda_2)\beta_2+\ldots+\overline{p}_1(\lambda_k)\beta_k=-\overline{p}_1(\lambda_{k+1})$$

. . . . . .

$$\overline{p}_{k-1}(\lambda_1)\beta_1+\overline{p}_{k-1}(\lambda_2)\beta_2+\ldots+\overline{p}_{k-1}(\lambda_k)\beta_k=-\overline{p}_{k-1}(\lambda_{k+1}).$$

In the second equation, $\overline{p}_1(\lambda)$ is a polynomial of degree 1; by adding λ_1 times the first equation to the second we may reduce the second to

$$\lambda_1\beta_1+\lambda_2\beta_2+\ldots+\lambda_k\beta_k=-\lambda_{k+1}.$$

In the third equation, $\overline{p}_2(\lambda)$ is a second-degree polynomial; a linear combination of the preceding two rows will reduce this equation to

$$\lambda_1^2\beta_1+\lambda_2^2\beta_2+\ldots+\lambda_k^2\beta_k=-\lambda_{k+1}^2.$$

Continuing in this way, we arrive at the equivalent system,

$$\beta_1+ \quad \beta_1+\ldots+ \quad \beta_k=-1$$

$$\lambda_1\beta_1+ \quad \lambda_2\beta_2+\ldots+ \quad \lambda_k\beta_k=-\lambda_{k+1}$$

. . . . . .

$$\lambda_1^{k-1}\beta_1+\lambda_2^{k-1}\beta_2+\ldots+\lambda_k^{k-1}\beta_k=-\lambda_{k+1}^{k-1}.$$

The solutions β_j are immediately expressible in terms of the function

$$V_k(b_1, b_2, \cdots, b_k)\equiv \prod_{\substack{l<m \\ l,\, m=1,2,\cdots,k}} (b_m-b_l).$$

From the identity

$$V_k(b_1, b_2, \cdots, b_k)=\begin{vmatrix} 1 & 1 & \cdots & 1 \\ b_1 & b_2 & \cdots & b_k \\ . & . & . & . \\ b_1^{k-1} & b_2^{k-1} & \cdots & b_k^{k-1} \end{vmatrix}$$

it follows that

$$\beta_1=(-1)^k\frac{V_k(\lambda_2, \cdots, \lambda_k, \lambda_{k+1})}{V_k(\lambda_1, \lambda_2, \cdots, \lambda_k)}$$

$$\beta_2=(-1)^{k-1}\frac{V_k(\lambda_1, \lambda_3, \cdots, \lambda_k, \lambda_{k+1})}{V_k(\lambda_1, \lambda_2, \cdots, \lambda_k)}$$

. . . . . .

$$\beta_k=-\frac{V_k(\lambda_1, \cdots, \lambda_{k-1}, \lambda_{k+1})}{V_k(\lambda_1, \lambda_2, \cdots, \lambda_k)}.$$

The L_j may now be computed from eq (40).

XI. Rate of Convergence

We shall establish some results on rate of convergence for the sequences of the previous section. The terminology is that of section 6. Recall the definition of δ_j, $j=1, 2, \ldots, r$, in eq (25). It is convenient to set

$$\delta_{r+1}=0.$$

We remark that for the lengths $|z_j^i|$ we have at once, from the corollary to theorem 10.1 and lemma 6.2, that $\{|z_j^i|\}$ is a sequence with ratio δ_j.

Theorem 11.1. For a fixed integer j, $j=2, 3, \ldots, r$, consider the *sequence* $\{\mu(z_j^i)\}$. (1) If $\delta_j^2>\delta_{j-1}\cdot\delta_{j+1}$, then $\{\lambda_j-\mu(z_j^i)\}$ is a positive, monotonic sequence with ratio $(\delta_j/\delta_{j-1})^2$. (2) If $\delta_j^2<\delta_{j-1}\cdot\delta_{j+1}$, then $\{\mu(z_j^i)-\lambda_j\}$ is a positive, monotonic sequence with ratio $(\delta_{j+1}/\delta_j)^2$.

We first assume $j<r$. Then from eq (36)

$$\mu(z_j^i)-\lambda_j=\frac{(z_j^i, Az_j^i)-\lambda_j|z_j^i|^2}{|z_j^i|^2}$$

$$=\frac{1}{|z_j^i|^2}\{(\lambda_1-\lambda_j)[p_{j-1}^i(\lambda_1)a_1^i]^2+\ldots$$

$$+(\lambda_{j-1}-\lambda_j)[p_{j-1}^i(\lambda_{j-1})a_{j-1}^i]^2$$

$$+(\lambda_{j+1}-\lambda_j)[p_{j-1}^i(\lambda_{j+1})a_{j+1}^i]^2+\ldots\}$$

$$=\frac{a_j^{i\,2}}{|z_j^i|^2}=\left\{b_1^i\frac{a_j^{i\,2}}{a_1^{i\,2}}+\ldots+b_{j-1}^i\frac{a_j^{i\,2}}{a_{j-1}^{i\,2}}\right.$$

$$\left.+b_{j+1}^i\frac{a_{j+1}^{i\,2}}{a_j^{i\,2}}+\ldots\right\}$$

where, by the corollary to theorem 10.1, the b_k^i have finite limits.

Now suppose $\delta_j^2>\delta_{j-1}\delta_{j+1}$. We write

$$(\mu(z_j^i)-\lambda_j)\frac{a_{j-1}^{i\,2}}{a_j^{i\,2}}=\frac{a_j^{i\,2}}{|z_j^i|^2}\left\{b_1^i\frac{a_{j-1}^{i\,2}}{a_1^{i\,2}}+\cdots+b_{j-2}^i\frac{a_{j-1}^{i\,2}}{a_{j-2}^{i\,2}}\right.$$

$$\left.+b_{j-1}^i+b_{j+1}^i\left(\frac{a_{j-1}^i a_{j+1}^i}{a_j^{i\,2}}\right)^2+\cdots\right\}. \qquad (41)$$

The factors of b_k^i (except for b_{j-1}^i) may be recognized as sequences of ratio less than 1, by lemma 6.1; hence these sequences tend to 0. Hence, by the above-mentioned corollary, the left side of eq (41) has the limit

$$-(\lambda_j-\lambda_{j-1})\frac{\overline{P}_{j-2}^2(\lambda_j)}{\overline{P}_{j-2}^2(\lambda_{j-1})}<0.$$

Using lemma 6.2 we deduce the first conclusion of the theorem.

Suppose next that $\delta_j^2<\delta_{j-1}\delta_{j+1}$. Then we write

$$(\mu(z_j^i)-\lambda_j)\frac{a_j^{i\,2}}{a_{j+1}^{i\,2}}=\frac{a_j^{i\,2}}{|z_j^i|^2}\left\{b_1^i\left(\frac{a_j^{i\,2}}{a_1^i a_{j+1}^i}\right)^2+\cdots\right.$$

$$\left.+b_{j-1}^i\left(\frac{a_j^{i\,2}}{a_{j-1}^i a_{j+1}^i}\right)^2+b_{j+1}^i+b_{j+2}^i\frac{a_{j+2}^{i\,2}}{a_{j+1}^{i\,2}}+\cdots\right\}.$$

All terms tend to 0 except b_{j+1}^i. Hence the left side has the limit

$$(\lambda_{j+1}-\lambda_j)\frac{\overline{p}_{j-1}^2(\lambda_{j+1})}{\overline{p}_{j-1}^2(\lambda_j)}>0.$$

From this we deduce the second conclusion of the theorem.

There remains the case $j=r$. In this instance only case (1) of the theorem is possible. The proof is identical with that given above for that case, except that no terms b_{j+1}^i and higher appear.

Theorem 11.1 leaves open the possibility of the equality $\delta_j^2=\delta_{j-1}\delta_{j+1}$. The likelihood of this condition holding in a numerical instance is very slight, but the problem is subtler than the instances of inequality and thereby has theoretical interest. We establish the following corollary.

Corollary. Suppose $\delta_j^2=\delta_{j-1}\delta_{j+1}$. Then

$$\mu(z_j^i)-\lambda_j=O(\delta_j^2/\delta_{j-1}^2).$$

To prove the corollary we require the following lemma.

Lemma 11.1. If $\delta_j^2=\delta_{j-1}\delta_{j+1}$ then the sequence

$$\left\{\frac{a_{j-1}^i a_{j+1}^i}{a_j^{i\,2}}\right\}$$

has a finite positive limit.

We remark that although the sequence in question has ratio 1 it does not follow from this that it has a limit.

By eq (16)

$$\frac{a_{j-1}^i a_{j+1}^i}{a_j^{i\,2}}=\frac{a_{j-1}^0 a_{j+1}^0}{a_j^{0\,2}}\cdot\prod_{k=0}^{i-1}\frac{\{1-\alpha(\lambda_{j-1}-\mu^k)\}\{1-\alpha(\lambda_{j+1}-\mu^k)\}}{\{1-\alpha(\lambda_j-\mu^k)\}^2} \tag{42}$$

where $\mu^k=\mu(z_1^k)$. The general term of the product may be written

$$\frac{[\delta_{j-1}+\alpha(\mu^k-\lambda_1)][\delta_{j+1}+\alpha(\mu^k-\lambda_1)]}{[\delta_j+\alpha(\mu^k-\lambda_1)]^2}.$$

Letting

$$\beta^k=\mu^k-\lambda_1,\quad K_1=\frac{\alpha}{\delta_{j-1}},\quad K_2=\frac{\alpha}{\delta_{j+1}},\quad K_3=\frac{\alpha}{\delta_j}$$

the general term becomes

$$\frac{(1+K_1\beta^k)(1+K_2\beta^k)}{(1+K_3\beta^k)^2}=1+b^k$$

where

$$b^k=\beta^k(K_1+K_2-2K_3)+\text{higher terms in }\beta^k.$$

The coefficient of β^k is positive. For,

$$K_1+K_2-2K_3=\frac{\alpha}{\delta_j}\left(\frac{\delta_j}{\delta_{j-1}}+\frac{\delta_j}{\delta_{j+1}}-2\right)$$

$$=\frac{\alpha}{\delta_j}\left(\rho+\frac{1}{\rho}-2\right)\qquad\left(\rho=\frac{\delta_j}{\delta_{j-1}}=\frac{\delta_{j+1}}{\delta_j}<1\right)$$

$$=\frac{\alpha\rho}{\delta_j}(\rho-1)^2>0.$$

It follows that for sufficiently large k, b^k is positive. Furthermore, by theorem 6.1 and lemma 6.2, b^k is a sequence with ratio δ_2^2. Hence $\sum_k b^k$ converges; thus the product $\Pi(1+b^k)$, and hence eq (42), converges to a positive limit as $i\to\infty$.

We turn to the proof of the corollary. As in the proof of the theorem we form eq (41). We see that all terms tend to 0 except the terms in b_{j-1}^i and b_{j+1}^i. The first of these has a finite limit, and, by lemma 11.1, the second does, too. This completes the proof.

Notice that for the corollary we cannot deduce that $\mu(z_j^i)-\lambda_j$ is a sequence with a ratio; this is because we do not know that the right side of eq (41) has a nonzero limit. In fact, it seems likely that this limit is zero.

Theorem 11.2. Let j be a fixed integer, $j=2, 3, \cdots, r$. Set $u_j^i=\frac{z_j^i}{|z_j^i|}$.

(1) If $\delta_j^2>\delta_{j-1}\delta_{j+1}$, then $|u_j^i-y_j|$ has ratio δ_j/δ_{j-1}.

(2) If $\delta_j^2<\delta_{j-1}\delta_{j+1}$, then $|u_j^i-y_j|$ has ratio δ_{j+1}/δ_j.

From the corollary to theorem 10.1 we see that

$$\frac{p_{j-1}^i(\lambda_j)\,a_j^i}{|z_j^i|}-1$$

tends to 0. We wish an estimate of the rate of convergence.

By eq (36),

$$|z_j^i|^2/[p_{j-1}^i(\lambda_j)\,a_j^i]^2-1=1/[p_{j-1}^i(\lambda_j)]^2\{[p_{j-1}^i(\lambda_1)\,a_1^i/a_j^i]^2+\cdots+[p_{j-1}^i(\lambda_{j-1})\,a_{j-1}^i/a_j^i]^2+[p_{j-1}^i(\lambda_{j+1})\,a_{j+1}^i/a_j^i]^2+\cdots\}.$$

Denote by c^{i2} the first term on the left. Under case (1) of the theorem we multiply both sides by

$$d_i^i=(a_{i-1}^i/a_i^i)^2.$$

As in the proof of theorem 11.1 we find that $(e^{i2}-1)d_i^i$ has a nonzero limit. Noting that $1-(1/e^i)=(e^{i2}-1)/e^i(1+e^i)$ and using $e^i\to 1$, we find that

$$\left(1-\frac{1}{e^i}\right)d_i^i$$

has a nonzero limit. (43)

On the other hand, under case (2) we find, by a similar argument that

$$\left(1-\frac{1}{e^i}\right)d_{i+1}^i \tag{44}$$

has a nonzero limit.
We have

$$|u_i^i-y_j|^2=\frac{1}{|z_j^i|^2}\left\{[p_{j-1}^i(\lambda_1)a_1^i]^2+\ \cdots\right.$$
$$+[p_{j-1}^i(\lambda_{j-1})a_{j-1}^i]^2+\left(\frac{1}{e^i}-1\right)^2$$
$$\left.+[p_{j-1}^i(\lambda_{j+1})a_{j+1}^i]^2+\ \cdots\right\}.$$

Consider condition (1) of the theorem. Multiply both sides of the above equation by d_i^i. From eq (43) and the corollary to theorem 10.1 we obtain that

$$|u_i^i-y_j|^2d_i^i \tag{45}$$

has a nonzero limit.
This proves the first part of the theorem. Consider next condition (2). This time we multiply both sides of the above equation by d_{i+1}^i. Using eq (44) we find

$$|u_i^i-y_j|d_{i+1}^i$$

has a nonzero limit. This completes the proof of the theorem.

Corollary. Suppose $\delta_i^2=\delta_{j-1}\delta_{j+1}$. Then

$$|u_i^i-y_j|=O(\delta_j/\delta_{j-1}).$$

Using lemma 11.1 we show as in the proof of the above theorem that the left side of eq (43) has a finite limit. Proceeding further as in the theorem we show that the left side of eq (45) also has a finite limit. This completes the argument.

The following result is of interest in connection with the preceding theorems.

Theorem 11.3. Suppose $r\geq 3$. Let j be an integer with $2\leq j\leq r-1$. If

$$\lambda_{j+1}-\lambda_j>\lambda_j-\lambda_{j-1},$$

then

$$\delta_i^2>\delta_{j-1}\delta_{j+1}.$$

Let $q_j=(\lambda_j-\lambda_1)/M$. From the definition of the δ_j

$$\delta_i^2-\delta_{j-1}\delta_{j+1}=(1-\beta q_j)^2-(1-\beta q_{j-1})(1-\beta q_{j+1})$$
$$=\beta^2[q_j^2-q_{j-1}q_{j+1}]+\beta D,$$

where

$$D=(q_{j+1}-q_j)-(q_j-q_{j-1})$$
$$=\frac{1}{M}[(\lambda_{j+1}-\lambda_j)-(\lambda_j-\lambda_{j-1})].$$

We find that

$$\delta_i^2-\delta_{j-1}\delta_{j+1}=\beta^2[(q_j-q_{j-1})^2-q_{j-1}D]+\beta D$$
$$=\beta^2(q_j-q_{j-1})^2+(\beta-\beta^2q_{j-1})D.$$

Since $0<\beta<1$, and $0\leq q_{j-1}<1$, it follows that the coefficient of D is positive. Hence the left side is positive if D is. This concludes the proof.

Notice that if $j=r$, the conclusion of the theorem holds, for, $\delta_{r+1}=0$. Also, if $r=2$, then $j=2$ is the only meaningful value of j and the conclusion likewise holds.

Theorem 9.2 enables us to prove an interesting result for the errors $\lambda_j-\mu_j^i$.

Theorem 11.4. Let

$$e_j^i=\lambda_j-\mu_j^i,\quad j=1,2,\ \ldots\ ,r;\quad i=0,1,2,\ \ldots\ .$$

Then for each i,

$$e_1^i+e_2^i+\ \ldots\ +e_j^i<0,\quad j=1,2,\ \ldots\ ,r-1,$$

$$e_1^i+e_2^i+\ \ldots\ +e_r^i=0.$$

Consider a fixed j, $j=1,2,\ \ldots\ ,r$, and a fixed i. As in theorem 9.2 let $\nu_1,\ \ldots\ ,\nu_j$ denote the roots of $p_j^i(\lambda)=0$. From the definition of the jth degree polynomial $p_j^i(\lambda)$ we can verify directly that the coefficient of λ^{j-1} is $-(\mu_1^i+\ \ldots\ +\mu_j^i)$. Hence

$$\sum_{k=1}^{j}\nu_k=\sum_{k=1}^{j}\mu_k^i.$$

From theorem 9.2,

$$\sum_{k=1}^{j}e_k^i=\sum_k\lambda_k-\sum_k\mu_k^i=\sum_k\lambda_k-\sum_k\nu_k=\sum_k(\lambda_k-\nu_k)\leq 0,$$

the equality holding just in case $j=r$.

XII. Concluding Remarks

Associated with a vector z_1 is the chain of subspaces

$$\mathscr{A}_1\subset\mathscr{A}_2\subset\cdots\subset\mathscr{A}_r$$

of section 9, together with the corresponding polynomials $p_j(\lambda)$, whose zeros give the characteristic roots of A relative to $\mathscr{A}_j$. Let $\nu_{1,j}$ denote a minimum solution of $p_j(\lambda)=0$. It is a consequence of theorem 9.1 that

$$\nu_{1,1} \geq \nu_{1,2} \geq \cdots \geq \nu_{1,r} = \lambda_1.$$

Similarly, if $\nu_{2,j}$ denotes the next smallest solution, then

$$\nu_{2,2} \geq \nu_{2,3} \geq \cdots \geq \nu_{2,r} = \lambda_2.$$

Analogous results hold for the higher characteristic roots. Thus, the characteristic roots of A can be found by successively finding the zeros of $p_2(\lambda)$, $p_3(\lambda)$, etc.[12] The procedure may be followed, for example, after a good approximation z_1^i has been obtained by, say, one of the gradient schemes.

The subspaces and their polynomials may be used in other ways in conjunction with the method of gradients. It was pointed out earlier that in the gradient methods we pass in each step from z_1 to a next approximation $\bar{z}_1$ which lies in the subspace $\mathscr{A}_2$ associated with z_1. One might consider choosing $\bar{z}_1$ in $\mathscr{A}_3$ or even a higher-dimensional subspace. The extra labor involved might be justified, or even essential, in an ill-conditioned problem in which several roots are clustered about λ_1 (or λ_r).

To illustrate, suppose that λ_1, λ_2, λ_3 are close together and reasonably isolated from the other roots. After a certain number of iterations it would be expected that the subspace $\mathscr{A}_3$ spanned by z_1, z_2, z_3 would contain only the characteristic vectors y_1, y_2, y_3 to a good approximation. At this stage, although the z's might not be good individual approximations to the y's, the characteristic vectors u_1, u_2, u_3 relative to $\mathscr{A}_3$ would provide such approximations. Further accuracy for y_1 would then be obtained by continuing the iteration process with u_1.

We proceed to derive the formulas for the characteristic roots $\nu_1 \leq \nu_2 \leq \nu_3$ and characteristic vectors u_1, u_2, u_3 relative to $\mathscr{A}_3$. The roots ν_k are the solutions of

$$p_3(\lambda) = (\lambda-\mu_1)(\lambda-\mu_2)(\lambda-\mu_3) - t_2^2(\lambda-\mu_3) - t_3^2(\lambda-\mu_1) = 0.$$

This cubic is most easily solved by introducing the new variable ρ and the constants μ, σ_k as follows:

$$\bar{\mu} = -(\mu_1+\mu_2+\mu_3), \quad \sigma_k = \mu_k - \bar{\mu}, \quad k = 1, 2, 3,$$

$$\lambda = \bar{\mu} + \rho.$$

The equation to be solved is now

$$\rho^3 + b\rho + c = 0,$$

where

$$b = \sigma_1\sigma_2 + \sigma_2\sigma_3 + \sigma_1\sigma_3 - t_2^2 - t_3^2, \quad c = t_2^2\sigma_3 + t_3^2\sigma_1 - \sigma_1\sigma_2\sigma_3.$$

The solutions may now be determined rapidly by, say, Newton's method. To find u_1, the characteristic vector relative to $\mathscr{A}_3$ belonging to ν_1, we write

$$u_1 = z_1 + \alpha z_2 + \beta z_3$$

[12] This is one of the methods developed by Lanczos.

and attempt to find α and β. To do this observe first that since u_k is a characteristic vector relative to $\mathscr{A}_3$ belonging to ν_k,

$$(u, Au_k - \nu_k u_k) = 0, \quad u \text{ in } \mathscr{A}_3, \quad k = 1, 2, 3. \tag{46}$$

Using eq (46) with $k=1$, and $u=z_1$, $u=z_3$ in turn we find, by eq (34),

$$0 = (\mu_1 - \nu_1)|z_1|^2 + \alpha|z_2|^2$$

$$0 = \alpha|z_3|^2 + \beta u_3|z_3|^2 - \beta\nu_1|z_3|^2.$$

Thus

$$\alpha = \frac{\nu_1 - \mu_1}{t_2^2}$$

$$\beta = \frac{\alpha}{\nu_1 - \mu_3},$$

so that

$$u_1 = z_1 + \frac{\nu_1 - \mu_1}{t_2^2}\left(z_2 + \frac{1}{\nu_1 - \mu_3} z_3\right).$$

The formulas for u_2 and u_3 are obtained by replacing ν_1 by ν_2 and ν_3, respectively.

Suppose the lower characteristic roots λ_1, λ_2, . . . and the higher characteristic roots λ_r, λ_{r-1}, . . . have been accurately calculated. The intermediate roots and vectors can then be calculated by the gradient procedure (12) as follows. We apply the procedure to an initial vector x^0, inducing (if necessary) and maintaining orthogonality to the vectors y_1, y_2, . . . and y_r, y_{r-1}, . . . by more or less frequent selection of γ^i as $1/(\lambda_j - \mu^i)$, with λ_j ranging over the known roots (see section 8). Use of the subspaces $\mathscr{A}_2$, $\mathscr{A}_3$, . . . may come into play as described above. Notice that as more roots are found the constant α in the method of fixed α may be chosen larger.

We remark that independent characteristic vectors belonging to a multiple characteristic root can be determined by varying the initial vector x^0.

We conclude by pointing out that the results of theorem 10.1 remain valid if the constant α satisfying eq (23) is replaced by the variable value of eq (28) where, in addition to eq (29) and (30), the condition

$$\frac{\beta_2}{\beta_1} < \min_{j=3,4,\ldots,r} \frac{\lambda_j - \lambda_1}{\lambda_{j-1} - \lambda_1}$$

is imposed. The purpose of this restriction is to guarantee that $a_k^i/a_j^i \to 0$ for $j<k$. The results on rate of convergence in section 11 may be modified to fit the new conditions.

Los Angeles, May 4, 1950.